



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 190359

TO: Celine Qian
Location: rem/2A64/2C70
Art Unit: 1636
Friday, June 09, 2006
Case Serial Number: 10/099663

From: Kristine Hensle
Location: Biotech-Chem Library
REM-1B69
Phone: (571) 272-4161

Kristine. Hensle@uspto.gov

Search Notes

Examiner Qian,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Kristine Hensle
Librarian (ASRC Aerospace)
STIC Biotech/Chem Library
(571)272-4161

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: May 28, 2006, 05:53:57 ; Search time 13877.8 Seconds
(without alignments)
10971.408 Million cell updates/sec

Title: US-10-099-663-1

Perfect score: 2381
Sequence: 1 agcttcctgcgcagaaaag.....atctgaagcctcattcag 2381

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl1.*
1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_dl:*
5: gb_pr:*
6: gb_ro:*
7: gb_rts:*
8: gb_sy:*
9: gb_un:*
10: gb_vl:*
11: gb_ov:*
12: gb_htg:*
13: gb_in:*
14: gb_om:*
15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	736.8	30.9	2847	11 AY254202	AY254202 Gallus ga
2	85	3.6	923	11 BX932340	BX932340 Gallus ga
3	81	3.4	899	11 BX930518	BX930518 Gallus ga
4	78.8	3.3	896	11 BX933597	BX933597 Gallus ga
5	64.2	2.7	1141	2 AR579680	AR579680 Sequence
6	64.2	2.7	1141	2 AR579680	AR579680 Sequence
7	62	2.6	3013	6 RATFABP	AX083744 Sequence
8	62	2.6	230897	12 AC130496	M18080 Rat Intesti
9	60.2	2.5	320940	12 AC152772	AC130496 Rattus no
10	57.8	2.4	1141	2 AR579680	AC152772 Bos tauri
11	57.8	2.4	1141	2 AX083744	AR579680 Sequence
12	57.6	2.4	320940	12 AC152772	AX083744 Sequence
13	56.6	2.4	5039	6 MUSFABP1	AC152772 Bos tauri
14	56.6	2.4	188707	12 AC158770	M65033 Mouse Fapb1
15	56.6	2.4	240767	6 AC155158	AC158770 Mus muscu
16	56.2	2.4	3509	14 AC155158	AC155158 Mus muscu
17	55	2.3	713	14 DQ126268	DQ126268 Sus scrofa
18	51.8	2.2	234081	13 AY911349	AY911349 Bos tauri
					PFMAL4P2

19	50.2	2.1	614	14 AY60624	AY60624 Sus scrofa
20	50.2	2.1	1522	13 AY701231	AY701231 Orconecte
21	50	2.1	564	6 RATFABP	M3592 Rat Intesti
22	50	2.1	187264	12 CR847945	CR847945 Danio rer
23	49.8	2.1	619	11 XELIFABP	L19946 Xenopus lae
24	49.8	2.1	5204	2 AY771590	AY771590 Sequence
25	49.8	2.1	5204	5 HOMFABP	M18079 Human, inte
26	49.8	2.1	200000	12 AC008077	AC008077 Homo sapi
27	49.2	2.1	426	2 CS177440	CS177440 Sequence
28	49.2	2.1	2000	2 AX655393	AX655393 Sequence
29	49	2.1	134124	6 AC154837	AC154837 Mus muscu
30	49	2.1	202804	6 AC021630	AC021630 Mus muscu
31	49	2.1	249943	13 AE014823	AE014823 Plasmodiu
32	48.6	2.0	176386	12 CR855271	CR855271 Danio rer
33	48.4	2.0	564	6 RATFABP1	K01180 Rat Intesti
34	48.4	2.0	832	2 AX972010	AX972010 Sequence
35	48.4	2.0	832	2 BD110729	BD110729 EST and e
36	48.4	2.0	832	2 AR415176	AR415176 Sequence
37	48.4	2.0	18876	13 CER3268	Z72509 Caenorhabdi
38	48.4	2.0	151090	11 AL929044	AL929044 Zebrafish
39	48.4	2.0	176472	11 CR762472	CR762472 Zebrafish
40	48.4	2.0	197110	5 AC104306	AC104306 Homo sapi
41	48.2	2.0	132254	13 AC116330	AC116330 Dictyoste
42	48.2	2.0	132254	13 AC116330	AC116330 Dictyoste
43	48	2.0	191840	5 AC092656	AC092656 Homo sapi
44	47.8	2.0	3853	11 AY266452	AY266452 Danio rer
45	47.8	2.0	110000	13 AC116984_1	Continuation (2 of

ALIGNMENTS

RESULT 1	AY254202	2847 bp	DNA	linear	VRT 22-APR-2003
LOCUS	AY254202				
DEFINITION	Gallus gallus intestinal fatty acid-binding protein gene, complete cds.				
ACCESSION	AY254202				
VERSION	AY254202.1	GI:30060211			
KEYWORDS					
SOURCE	Gallus gallus (chicken)				
ORGANISM	Gallus gallus				
REFERENCE	1 (bases 1 to 2847)				
AUTHORS	Wang, Q., Li, H., Wang, Y. and Zhao, J.				
TITLE	Cloning and characterization of chicken I-FABP gene				
JOURNAL	Unpublished				
AUTHORS	2 (bases 1 to 2847)				
TITLE	Wang, Q., Li, H., Wang, Y. and Zhao, J.				
JOURNAL	Submitted (11-MAR-2003) Animal Science & Technology, Northeast Agricultural University, 59 Mucal Street, Harbin, Heilongjiang 150030, China				
FEATURES	Location/Qualifiers				
source	1..2847				
	/organism="Gallus gallus"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9031"				
	/rfeature_type="blood"				
	/note="Breed: Arbor Acres broiler"				
mRNA	join(21..87,720..892,1465..1572,2416..2466)				
	/product="intestinal fatty acid-binding protein"				
	join(21..87,720..892,1465..1572,2416..2466)				
	/note="FABP"				
CDS	/codon_start=1				
	/product="intestinal fatty acid-binding protein"				
	/protein_id="AAP101.1"				
	/db_xref="GI:30060212"				
	/translation="MAENGWTKIEKKNENYKPEAKGVNVMKRLGADHDUKLTIOODGNKFLVSSNFRITDIEFTLVGSFVSLADGTELSGNWLEGNKVLGTRIDKNGRYLTAVREIVGSELIQTIVYEGVEAKRIFKKE"				

ORIGIN

Query Match 30.9%; Score 736.8; DB 11; Length 2847;
Best Local Similarity 98.1%; Pred. No. 8.9e-221;
Matches 751; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

QY 1613 CTACAGACAGAAAGATGGCATTTAAGGTAAGTGAAGAAATGAGAACTG 1672
DB 7 CGAGAGACAGAAAGATGGCATTTAAGGTAAGTGAAGAAATGAGAACTG 66
QY 1673 AAAAATTCATGAGAGCAATGGGTAAGCTTTTGAATGCCCTTCTTAAAGCAGAT 1732
DB 67 AAAAATTCATGAGAGCAATGGGTAAGCTTTTGAATGCCCTTCTTAAAGCAGAT 126
QY 1733 ACCACTAGCGCGGAAATACAACTTAAGCTGTTCAAGAACTTACATCTGCTAAGCTGTC 1792
DB 127 ACCACTAGCGCGGAAATACAACTTAAGCTGTTCAAGAACTTACATCTGCTAAGCTGTC 186
QY 1793 TTTGTTGCTGCTATTTTGCCTTGACATTCGCCCTGACCTTATTTTGAAGACTCTAT 1852
DB 187 TTTGTTGCTGCTATTTTGCCTTGACATTCGCCCTGACCTTATTTTGAAGACTCTAT 246
QY 1853 AAGGGGAAATACAGAGAAAGAAACATTTGATTTTATTTGATTCGATATCTTATGC 1912
DB 247 AAGGGGAAATACAGAGAAAGAAACATTTGATTTTATTTGATTCGATATCTTATGC 306
QY 1913 ATTAGCTAATTCAGTAGAGGCAATTCAGAGAAATTTAAATGAAATTAATGTAAGCA 1972
DB 307 ATTAGCTAATTCAGTAGAGGCAATTCAGAGAAATTTAAATGAAATTAATGTAAGCA 366
QY 1973 ATATTATTTTGAATGAAGCTGTTGAAATTAACAGAGAGGAAATGCTGCTCCAGT 2032
DB 367 ATATTATTTTGAATGAAGCTGTTGAAATTAACAGAGAGGAAATGCTGCTCCAGT 426
QY 2033 TTTGAGAGACACATGATTTT---GAGTCATTTTAACTGCTAGGCTTACTTTAAGCTT 2089
DB 427 TTTGAGAGACACATGATTTTGAATTTTAACTGCTAGGCTTACTTTAAGCTT 486
QY 2089 GTACAAATCTGCTGTAATATGATGTAATTAACATTAATCTTAAATGTAAGTGTGTA 2149
DB 487 GTACAAATCTGCTGTAATATGATGTAATTAACATTAATCTTAAATGTAAGTGTGTA 546
QY 2150 TTACAGGCTGAACCTGCTCAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2209
DB 547 TTACAGGCTGAACCTGCTCAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 606
QY 2210 CTGGGCTAAGCTCCCTCAACTACAGAAAGGATCAATTAATAAATGCAATGATGTTCT 2269
DB 607 CTGGGCTAAGCTCCCTCAACTACAGAAAGGATCAATTAATAAATGCAATGATGTTCT 666
QY 2270 ATTTGTTTTTCTCTGCTGATGTAATGATTTATTTATTTTATTTTATTTTATTTAGCGGTGA 2329
DB 667 ATTTGTTTTTCTCTGCTGATGTAATGATTTATTTATTTATTTTATTTTATTTTATTTAGCGGTGA 726
QY 2330 TGTGATGAAAGAAAGTTAGAGCGCCAGATTAATCTGAAGCTCACTATTCAG 2381
DB 727 TGTGATGAAAGAAAGTTAGAGCGCCAGATTAATCTGAAGCTCACTATTCAG 778

RESULT 2
LOCUS BX932340 923 bp mRNA linear VRT 30-MAR-2004
DEFINITION Gallus gallus finished cDNA, clone CHEST35022.
ACCESSION BX932340
VERSION BX932340.2 GI:46017395
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 923)
AUTHORS Boardman, P.E., Bonfield, J.K., Brown, W.R.A., Carder, C., Chalk, S.E.,

TITLE
JOURNAL
COMMENT

FEATURES
source

ORIGIN

Query Match 3.6%; Score 85; DB 11; Length 923;
Best Local Similarity 100.0%; Pred. No. 1e-14;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1610 TGGCTTCAAGACAGAAAGATGGCAATTTAAGGTAAGTGAAGAAATGAGAACT 1669
DB 1 TGGCTTCAAGACAGAAAGATGGCAATTTAAGGTAAGTGAAGAAATGAGAACT 60
QY 1670 ATGAAATATTCATGAGAGCAATGGG 1694
DB 61 ATGAAATATTCATGAGAGCAATGGG 85

RESULT 3
LOCUS BX930518 899 bp mRNA linear VRT 02-FEB-2004
DEFINITION Gallus gallus finished cDNA, clone CHEST640b17.
ACCESSION BX930518
VERSION BX930518.1 GI:41631046
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 899)
AUTHORS Boardman, P.E., Bonfield, J.K., Brown, W.R.A., Carder, C., Chalk, S.E.,
Croning, M.D.R., Davies, R.M., Francis, M.D., Grafham, D.V.,
Hubbard, S.J., Humphray, S.J., Hunt, P.J., Maddison, M., McLaren, S.R.,
Niblett, D., Overton, I.M., Rogers, J., Scott, C.E., Taylor, R.G.,
Tickle, C. and Wilson, S.A.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-2004) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: chickes@dm.sanger.ac.uk
COMMENT BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
sequencing project.
This sequence is from the
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
from a library constructed by Elizabeth Bosch. cDNA was prepared
from RNA extracted from heart, normalised, and poly A-tailed.
EcoRI-NotI cut cDNA was then ligated into the vector. Vector:
pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
coli DH10B.
location/Qualifiers
1..923
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST35022"
/clone_11b="CSBCHN70"
/dev_stage="stage 36"

JOURNAL Unpublished (1987)
 REFERENCE 2 (bases 917 to 1217)
 AUTHORS Sweetser,D.A., Birkenmeier,E.H., Kliaek,I.J., Zollman,S.,
 Sparke,R.S., Mohandes,T., Lusis,A.J. and Gordon,J.I.
 TITLE The human and rodent intestinal fatty acid binding protein genes. A
 comparative analysis of their structure, expression, and linkage
 relationships
 JOURNAL J. Biol. Chem. 262 (33), 16060-16071 (1987)
 PUBMED 2824476
 COMMENT Original source text: Rat (Sprague Dawley) intestinal epithelial
 lining DNA, clone lambda-RIRABP.
 Draft entry and computer-readable sequence for [2] kindly provided
 by D.Sweetser, 19-JAN-1988.
 FEATURES
 source location/Qualifiers
 1..3013
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 repeat_region 195..406
 /note="B2 repeat"
 prim_transcript 1155..>3013
 /note="FABP mRNA (alt.) and introns"
 prim_transcript 1179..>3013
 /note="FABP mRNA (alt.) and introns"
 join(1215..1281,2541..2713)
 CDS
 /note="fatty acid binding protein"
 /codon_start=1
 /protein_id="AA041133.1"
 /db_xref="GI:204072"
 /translation="MAFDGTMKVDNENYKFMKMGINVVKKAGANDLKLTTQEG
 GNKTVKSSNFRNIDVVEELGVPAVSLADGTBLT"
 <1215..1281
 /note="fatty acid binding protein"
 /number=1
 intron 1282..2540
 /note="FABP intron A"
 2541..>2713
 /note="fatty acid binding protein"
 /number=2
 intron 2714..>3013
 /note="FABP intron B"
 ORIGIN 1 bp upstream of EcoRI site.
 Query Match 2.6% Score 62; DB 6; Length 3013;
 Best Local Similarity 59.3%; Pred. No. 1.9e-07;
 Matches 127; Conservative 0; Mismatches 80; Indels 7; Gaps 1;
 QY 1515 TTGAGCTTTCACGACATCATCATGTTGCTTCTGATTAAGCTGTTGATTA
 1574
 DB 1096 TTGAACCTTGAACCTTCACATCATGATGATGTTGCGAAGATTAAGAAATTA
 1155
 QY 1575 TTCTCTTTCGAAAGCTGCTGCTACTTACGAAAGTC-----TGCTTCACAGAGAAGA
 1627
 DB 1156 TTCTCTTTCGAGACAGACGAACTCTGCTTTCCTGAGGACACACAGCTGACATCA
 1215
 QY 1628 TGGCATTTAAGCTTTCGAAATAATGAGAAATGAGAAATTCATGAGAG 1687
 DB 1216 TGGCATTTGATGACCTTGAAGAGTACCGGAATGAGAACTTATTAAGTTTCA
 1275
 QY 1688 CAATGGGTAGCCTTACTTTTGAAGTCTTCT 1721
 DB 1276 AATGGGTAGGCGCTGCTCTGCTATTGCT 1309

ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Rattus.
 REFERENCE
 1 (bases 1 to 230897)
 Muzny,D.,Marle., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Albrooks,S., Amth,A., Angiano,D.,
 Anyalebechi,V., Ayodeji,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Butrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Dermo,C., Ding,Y., Dinh,H., Divya,K.,
 Duper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebregeorgis,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guetara,W.,
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,W.,
 Hollins,B., Howells,S., Huik,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolyet,A.,
 Karpathy,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorenshaw,L., Louised,H., Lozada,R.J., Lu,K., Ma,J.,
 Maneshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
 Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Mundaasa,M., Murphy,M., Natl,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwokoelamleh,O., Okwou,G., Olarnuagsoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
 Plopper,F., Poindecker,A., Popovic,D., Prims,E., Pu,L.,
 Puzo,M., Quiroz,J., Rachlin,B., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Slason,I., Sitter,C.D., Smajz,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Tjorjo,Z., Umami,K.,
 Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wleczky,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinsteck,G. and Gibbs,R.A.
 TITLE
 JOURNAL
 REFERENCE
 2 (bases 1 to 230897)
 Worley,K.C.
 TITLE
 JOURNAL
 REFERENCE
 3 (bases 1 to 230897)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (09-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 9, 2002 this sequence version replaced gi:23101530.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas

COMMENT

of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jan 23, 2005 this sequence version replaced gi:55819826.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: PATP

Center clone name: CH240-10M22

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 288280 bases at least Q40

Consensus quality: 294971 bases at least Q30

Consensus quality: 301063 bases at least Q20

Estimated insert size: 292138; sum-of-contigs estimation

Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbankdraft_data.html)

NOTE: This sequence may represent more than one clone.

NOTE: This is a 'working draft' sequence. It currently

consists of 81 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 5541: contig of 5541 bp in length
5542 5591: gap of 50 bp
5592 9604: contig of 4013 bp in length
9605 9704: gap of unknown length
9705 11503: contig of 1799 bp in length
11504 11802: gap of 299 bp
11803 20154: contig of 8352 bp in length
20155 20254: gap of unknown length
20255 24517: contig of 4263 bp in length
24518 24567: gap of 50 bp
24568 25759: contig of 1192 bp in length
25760 26270: gap of 511 bp
26271 29359: contig of 3089 bp in length
29360 32345: gap of unknown length
32346 32395: contig of 2886 bp in length
32396 38577: gap of 50 bp
38578 38677: contig of 6182 bp in length
38679 40965: gap of unknown length
40966 41015: contig of 2288 bp in length
41016 47383: contig of 6368 bp in length
47384 47483: gap of unknown length
47484 48761: contig of 1278 bp in length
48762 48842: gap of 81 bp
48843 51170: contig of 2328 bp in length
51171 51270: gap of unknown length
51271 53092: contig of 1822 bp in length
53093 53142: gap of 50 bp
53143 55224: contig of 2082 bp in length
55225 55313: gap of 89 bp
55314 57275: contig of 1962 bp in length

57276 57375: gap of unknown length
57376 63877: contig of 6502 bp in length
63878 63927: gap of 50 bp
63928 65212: contig of 1285 bp in length
65213 65262: gap of 50 bp
65263 66517: contig of 1555 bp in length
66518 66617: gap of unknown length
66619 70198: contig of 3581 bp in length
70199 71339: gap of 1141 bp
71340 75208: contig of 3669 bp in length
75209 75308: gap of unknown length
75309 97232: contig of 21924 bp in length
97233 97282: gap of 50 bp
97283 100331: contig of 3949 bp in length
100332 101076: gap of 745 bp
101077 102495: contig of 1419 bp in length
102496 102595: gap of unknown length
102596 104484: contig of 1889 bp in length
104485 104708: gap of 224 bp
104709 109824: contig of 5116 bp in length
109825 109925: gap of unknown length
109926 115197: contig of 5273 bp in length
115198 115247: gap of 50 bp
115248 123071: contig of 7824 bp in length
123072 123171: gap of unknown length
123172 131072: contig of 7901 bp in length
131073 131122: gap of 50 bp
131123 135372: contig of 4250 bp in length
135373 135472: gap of unknown length
135473 140453: contig of 4861 bp in length
140454 140503: gap of 50 bp
140504 143862: contig of 3359 bp in length
143863 143962: gap of unknown length
143963 145121: contig of 1159 bp in length
145122 145518: gap of 397 bp
145519 147216: contig of 1298 bp in length
147217 147316: gap of unknown length
147317 149768: contig of 2452 bp in length
149769 149818: gap of 50 bp
149819 152644: contig of 2826 bp in length
152645 152694: gap of 50 bp
152695 158870: contig of 6176 bp in length
158871 159535: gap of 665 bp
159536 160976: contig of 1441 bp in length
160977 162113: gap of 1137 bp
162114 163749: contig of 1636 bp in length
163750 163849: gap of unknown length
163850 169431: contig of 5582 bp in length
169432 169481: gap of 50 bp
169482 172813: contig of 3332 bp in length
172814 172913: gap of unknown length
172914 177020: contig of 4107 bp in length
177021 177070: gap of 50 bp
177071 179744: contig of 2674 bp in length
179745 179912: gap of 168 bp
179913 182027: contig of 2115 bp in length
182028 182077: gap of 50 bp
182078 187508: contig of 5431 bp in length
187509 187558: gap of 50 bp
187559 187559: gap of 50 bp
187560 196517: contig of 8959 bp in length
196518 196567: gap of 50 bp
196568 200212: contig of 3645 bp in length
200213 200410: gap of 198 bp

Query Match 2.5%; Score 60.2; DB 12; Length 320940;
Best Local Similarity 59.4%; Pred. No. 7.1e-07;
Matches 142; Conservative 0; Mismatches 88; Indels 9; Gaps 2;

QY 1492 CCTGACAGATGACGACAGTTCAGCTTACCGCAGCAGCAGTTCAGT-AAAATGC 1550
DB 162201 CATTAATGAGTCAAGAGGCGCTTGAACCTTAATCACTTCACATCAATATGAGCTGG 162260
QY 1551 TTCTCTATAGCCCTGTCATTAATTCCTTTGCAAGCTCTGCTACTTACGAGA----- 1605

Db	Accession	Score	DB 2	Length	1141
Db	162261	TTCTAAGATTAAGAAAATAGAAATTAATTCCTCTCTAAGAGACAGACTGAATCTCGATTGCT	162320		
Qy	1606	---AGCTGCTCCTACAGACAGAAAGATGCGATTTAACGGTACTTGGAAAATATGAGAAAAT	1662		
Db	162321	CAGAGCGCTGAATTAACACAGAGACATAGCGGCTTTGATGTAATTGGAAGGAGACAGAAAT	162380		
Qy	1663	GAGAACTATGAAAAATTCATGGAAGAAATGGGTAAGCCTACTTTTGAATCCCTCT	1721		
Db	162381	GAGAACTATGAAAAATTCATGGAAGAAATGGGTAAGCCTACTTTAATCAAGCCTGT	162439		
RESULT 10					
AR579680	AR579680	1141 bp	DNA	linear	PAT 14-DEC-2004
LOCUS	Sequence	22	from patent US 6784342.		
DEFINITION	AR579680				
ACCESSION	AR579680.1	GI:56583130			
VERSION					
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 1141)				
TITLE	Kunet, L. and Clemens, S.				
JOURNAL	Regulation of embryonic transcription in plants				
	Patent: US 6784342-A 22 31-AUG-2004,				
	The University of British Columbia; Vancouver;				
	CAX;				
FEATURES	location/Qualifiers				
source	1..1141				
	/organism="unknown"				
	/mol_type="genomic DNA"				
ORIGIN					
Query Match	2.4%; Score 57.8; DB 2; Length 1141;				
Best Local Similarity	10.9%; Pred. No. 4e-06;				
Matches	82; Conservative 263; Mismatches 407; Indels 0; Gaps 0;				
Qy	1565	TGTTGATTAATTCCTCTTGGCAAAGCTCTGCTACTTACAGAAAGCTGCTACAGACAGAA	1624		
Db	201	DRDMSBRKMYGMBMWKMSYDYVYVWWDMDCKRKRVRVTRGRKNTVWMBTMR	260		
Qy	1625	AGATGCGATTTACGGTACTTGAAAAATAGAAAAATAGAACTATGAAAAATTCATGG	1684		
Db	261	RRYNGTBMAYRRYRTVNNNNNNNAKMKCKRAXWGMRAVNSCTTWSKTRKVRTSC	320		
Qy	1685	AAGCAATGGGTAAAGCTTACTTTTGAATGCTCTTAAAGCAGATACCACTACGGGG	1744		
Db	321	WANNCRAGDAKDKHKWMSAAMGYVNNNNNNNTYKKARIBABWDVWMSAMKWHAN	380		
Qy	1745	GAATACAACTTAAGCTGTGATGAATCACTGCTGCTAAGCTGCTCTTGTGTGCTG	1804		
Db	381	AAHSRRKKWTYKRTVNNNNNGTTMKRMAYMKMDMDGTYNNNNNGRTYYGWTK	440		
Qy	1805	TATTTGGCCTTGACATGCGCTGCACTTAATTTGAAAAAGCTATAGAGGGAATAC	1864		
Db	441	NKQWVYKWKANCKRPMDHKTCTHNNTTWKQKTYNNNCYKWSMTGKSHRBAAY	500		
Qy	1865	AAGGAAGAAAACATTCGATTTTATGATTCATGGCATATCTTAAGCTTAAT	1924		
Db	501	TWYMMWRRYVHAANNNDYMKACTWYBYCSKMMNNYAAWYTKSSWNTSRYYRMKT	560		
Qy	1925	CCAGTAGAGGCAATCCAGAGAAATTTAATGAATATATGTAAGAAATATATTTTGA	1984		
Db	621	HMKMYAKKYAAGGSGNNNNNNNNNNNNNNNNATCAADDYTAASRTYMAAAKYYK	680		
Qy	1985	TAAAGCTGTTGAAAATTAACAAGAGGAATGCTGCTCCAGTTTGGAGAACAC	2044		
Db	2045	ACATATTTGAAGTCAATTTTAAAGATGCTAGCTTACTTTAAGCTTGACAACTGCTGT	2104		
Qy	661	BAANNAYTTHANNWGCANNAPTBRTTWKNNNNNNAGTAKNNNNNNNAKVASAAKYAAA	740		

QY	2105	AATATGATGTAAACATACTATCTCTAGTGTAGTACTGTGTATTTGATTTACAGGCTGAACAC	2164
Db	741	AAVAAKHKRRMANKMAMRGMBADAAABTTDDKKNAGAYTKTITTTTTNNNTYRGRAYTTAARD	800
QY	2165	TGCGCTCAGTGAAGGATGAGAGAGTAAAGACTCTGAGTCAGAAATCTCGGCGTAAGCTCCC	2224
Db	801	GAANNNNNNNNNNNNNNNGWSDMWVYVWMAVANYGTNNNNNNNNNNNAVAMWTKMYTTDDR	860
QY	2225	TCAACTACAGAAAAAGTCAACATAAATAATGCAACATGATGTCTCATATTTGTCTTCTCT	2284
Db	861	WRBAYTNNNNNNNNNNMAYYGAYADYATYMSDTCDAAMKWDATNNATYINRGTAIRKTTNN	920
QY	2285	GCTTGATGTAAATTGATTTATTTATTTT	2316
Db	921	NNNTYTKTYBHAAMNNNNNNCKMCTATHTW	952
RESULT	11		
LOCUS	AX083744	1141 bp	DNA linear PAT 28-FEB-2001
DEFINITION	Sequence 22 from Patent WO0111061.		
ACCESSION	AX083744		
VERSION	AX083744.1	GI:13185472	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS	Kunst, L. and Clemens, S.		
TITLE	Regulation of embryonic transcription in plants		
JOURNAL	Patent: WO 011061-A 22 15-FEB-2001; UNIVERSITY OF BRITISH COLUMBIA (CA)		
FEATURES			
Source	1. 1141 /organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630"		
	1. 1141 /note="consensus sequence of A.t., L.a., and B.n. FAD1 promoters"		
promoter			
Query Match	2.4%: Score 57.8; DB 2; Length 1141;		
Best Local Similarity	10.9%: Pred. No. 4e-06;		
Matches	82; Conservative 263; Mismatches 407; Indels 0; Gaps 0;		
QY	1565	TGTTATATAATCTCTTTGCAAGCTCTGTACTTACCAAGATCTGCTACAGACAGAA	1624
Db	201	DRDMSBKMMYGMWBWKKMSYDVYVWVMDMCKRVRVRVTRGMRMYAMVTAHR	260
QY	1625	AGATGGCATTTAAGCGTACTTGAAATAAGAAATAAGAACTATGAAATAATTCATCG	1684
Db	261	RRYNNGMTBAAYRRVTWNNNNNNNAKAMCKAKTYGMRABVNSTCTTWSKTKVATSC	320
QY	1685	AAGCAATGGGTAAAGCTACTTTTGTGAATGCTTTTAAAGCAGATACACATACGGC	1744
Db	321	WANNCRADGANRDKHMKWMSAMGYVNNNNNNMMTYKKAHBAKMDVMSAKMKWHAN	380
QY	1745	GAATACAACTTAAGCTGTCTATGAACTAACAATCTGGCTAAGCTGCTTGTGTCTGC	1804
Db	381	AAHYSRKMTRBKRTMVNNNNNGTTMWMKMAHYKMKMDMBGTNNNNNGGRITYGWTX	440
QY	1805	TATTTGACCTTGACATTTGCGCTGCACTTATTTTGAAGAAGCTCTATAGAGGGAATAC	1866
Db	441	NKKMTYYKKKXANNCKMRAMHKTCVHNNTTWMKKTYMNNCKYSMSMTNGSKSHRAAIV	500
QY	1865	AAGGAAGAAAACATTCGATTTTATTTGACATGGGATAATCTTAGCATTTAGCTAAT	1924
Db	501	TYTMMWREYHAHANNNDMMYWKACTWYKVBVCSKMMNNYAAMYTKSSMNTSRYYRKT	560
QY	1925	CCAGTAGAGGCAATTCAGAGAAATTTAAATGAATTTATGTGAAGAAATATATTTTGA	1988

[illegible]

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Szatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, F., Telitola, B., Thomas, N., Thomas, S., Umanan, R., Vasquez, L., Vera, V., Verdusco, D., Villalon, D., Vinson, R., Waldron, L.M., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Xu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 320940)
Worley, K.C.

Submitted (18-NOV-2004) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 320940)
Worley, K.C.

Direct Submission
Submitted (23-JAN-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jan 23, 2005 this sequence version replaced gi:55819826.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly ('a' 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information -----
Center project name: FARP
Center clone name: CH240-10M22

----- Summary Statistics -----
Assembly program: Atlas 3.0;
Consensus quality: 288280 bases at least Q40
Consensus quality: 294971 bases at least Q30
Consensus quality: 301063 bases at least Q20
Estimated insert size: 292136; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* consists of 81 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 5541: contig of 5541 bp in length
* 5591: gap of 50 bp
* 5592 9604: contig of 4013 bp in length
* 9605 9704: gap of unknown length
* 9705 11503: contig of 1799 bp in length
* 11504 11802: gap of 299 bp
* 11803 20154: contig of 8352 bp in length
* 20155 20254: gap of unknown length
* 20255 24517: contig of 4263 bp in length
* 24518 24567: gap of 50 bp
* 25759: contig of 1192 bp in length
* 25759: contig of 1192 bp in length

25760	26270:	gap of 511 bp
26271	29359:	contig of 3089 bp in length
29360	29359:	gap of unknown length
29360	32345:	contig of 2886 bp in length
32346	32395:	gap of 50 bp
32396	38577:	contig of 6182 bp in length
38578	38577:	gap of unknown length
38578	40965:	contig of 2288 bp in length
40966	41013:	gap of 50 bp
40966	47383:	contig of 6368 bp in length
47384	47383:	gap of unknown length
47384	48761:	contig of 1278 bp in length
48762	48842:	gap of 81 bp
48843	51170:	contig of 2328 bp in length
51171	51270:	gap of unknown length
51271	53092:	contig of 1822 bp in length
53093	55314:	gap of 50 bp
55313	55224:	contig of 2082 bp in length
55225	55313:	gap of 89 bp
55314	55275:	contig of 1962 bp in length
55276	57375:	gap of unknown length
57376	63877:	contig of 6502 bp in length
63878	63927:	gap of 50 bp
63928	65212:	contig of 1285 bp in length
65213	65262:	gap of 50 bp
65263	66517:	contig of 1255 bp in length
66518	66617:	gap of unknown length
66618	70199:	contig of 3581 bp in length
70199	71339:	gap of 1141 bp
71340	75208:	contig of 3869 bp in length
75209	97232:	gap of unknown length
97233	97232:	contig of 21924 bp in length
97233	97282:	gap of 50 bp
97283	100331:	contig of 3049 bp in length
100332	101076:	gap of 745 bp
101077	102495:	contig of 1419 bp in length
102496	102595:	gap of unknown length
102596	104484:	contig of 1869 bp in length
104485	104708:	gap of 224 bp
104709	109824:	contig of 5116 bp in length
109825	109924:	gap of unknown length
109925	115197:	contig of 5273 bp in length
115198	115247:	gap of 50 bp
115248	123071:	contig of 7824 bp in length
123072	123171:	gap of unknown length
123172	131072:	contig of 7901 bp in length
131073	131122:	gap of 50 bp
131123	135472:	contig of 4250 bp in length
135473	135472:	gap of unknown length
135473	140453:	contig of 4981 bp in length
140454	140503:	gap of 50 bp
140504	143862:	contig of 3359 bp in length
143863	143962:	gap of unknown length
143963	145121:	contig of 1159 bp in length
145122	145518:	gap of 397 bp
145519	147216:	contig of 1698 bp in length
147217	147316:	gap of unknown length
147317	149768:	contig of 2452 bp in length
149769	149818:	gap of 50 bp
149819	152644:	contig of 2826 bp in length
152645	152694:	gap of 50 bp
152695	158870:	contig of 6176 bp in length
158871	158935:	gap of 665 bp
158936	160976:	contig of 1441 bp in length
160977	162113:	gap of 1137 bp
162114	163749:	contig of 1636 bp in length
163750	163849:	gap of unknown length
163850	169431:	contig of 5582 bp in length
169432	169481:	gap of 50 bp
169482	172813:	contig of 3332 bp in length
172814	172913:	gap of unknown length
172914	177020:	contig of 4107 bp in length
177021	177070:	gap of 50 bp

Query Match	2.44;	Score 57.6;	DB 12;	Length 320940;
Beet local Similarity	65.64;	Pred. No. 4.7e-06;		
Matches	84;	Conservative	44;	Indels 0;
		Mismatches	44;	Gaps 0;

1594 CTACTTACCAAGTCTGCGCTACAGACAGAAGATGGCATTTAA CGGTACTTGAAAAATA 1653
| | | | | | | | | | | | | | | | | |
22649 CTGAATTGCTCAGAGGCTGTAATTAACCAAGACCATTGGCGTTTAATGGTA CTTGGAAGTTA 2285

1654 GAGAAAATGAACTATGAAAAATTCATGGAGCAATGGTAAAGCTTACTTTTTGAA 1713
 22589 GACGAAATGGAACCTTGAAAGTTCATGGAAAAAAGGTAAAGTCTTAATACTTAA 225

Qy 1714 TGCCTTC 1721
|||
||
22529 CAGCTTGT 22522
bb

RESULT 13

Gene	Size (bp)	Structure	Accession
MUSFAB1 LOCUS	5039 bp	DNA	ROD 27-APR-1999
MUSFAB1		linear	
Mouse Fabp1 gene, exons 1-4.			

DATE	1993-03-10
ACCESSION	M65033
VERSION	M65033.1
GI	193218

KEYWORDS
SOURCE
Mus musculus (house mouse)
Mus musculus
Mus musculus

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sclerogomphi; Muridae; Murinae; Mus.

REFERENCE
1 (sites)
Green, R.P., Cohn, S.M., Sacchetti, J.C., Jackson, K.E. and

TITLE Gordon, J. I.
The mouse intestinal fatty acid binding protein gene: nucleotide sequence, pattern of developmental and regional expression, and

proposed structure of its protein product
DNA Cell Biol. 11 (1), 31-41 (1992)

PUBLISHED
1/3/54:55
REFERENCE
2 (pages 1 to 5039)
AUTHORS
Green, R. P., Sacchetti, J. C., Jackson, K. E., Cohn, S. M. and

TITLE Gordon, J.I.
The mouse fatty acid binding protein gene (Fabp1): nucleotide sequence and structure of its protein

sequence analysis and predictions of the structure of the product (1991)

FRATRES	JOURNAL	COMMENT	location/Qualifiers
unpublished (1992)	Original	source text: Mus musculus (strain DBA/2J)	Adult Liver DNA

1. .5039
/organism="Mus musculus"

```

/mol_type="genomic DNA"
/strain="DBA/2J"
/4t_mol_fraction=10000

```

```

/od_xref=caadn:lvv30
/cissue_type="Liver"
/dev_stage="Adult"

```

```
misc_feature
467..547
/note="Domain I: homology with human I-EABP promoter
```

repeat_region	region'; putative"
524..537	

/note="One of two repeats homologous to the one found in number of other lipid binding proteins which are expressed in small intestine"; putative"

misc_feature
727..817
/note="Domain II: homology to human I-FABP promoter"


```

misc_feature      region: putative"
                   870..966
                   /note="Domain III: homology with human I-FABP promoter
                   region: putative"
repeat_region     884..897
                   /note="Second of two repeats homologous to consensus
                   repeat: putative"
                   937..942
TATA_signal       /note="putative"
                   967..1069
exon              /label=Exon1
                   2326..2497
exon              /note="putative"
                   /label=Exon2
                   4291..4398
exon              /note="putative"
                   /label=Exon3
                   4818..5039
exon              /note="putative"
                   /label=Exon4

```

ORIGIN

```

Query Match      2.4%; Score 56.6; DB 6; Length 5039;
Best Local Similarity 58.5%; Pred. No. 9.6e-06;
Matches 121; Conservative 0; Mismatches 79; Indels 7; Gaps 1;

QY 1516 TGAGCTTTAGCAGCAGCATATCATATGTAATTGCTTCCTGATAGCCTGTTATAT 1575
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 884 TGAACCTTGCAACTTCCACATATCAAGTAATGTTGTAAGATTAAGAAATTAAT 943
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1576 TCTCTTTGCAAGCTCTGCTACTTACCAAGATC-----TGCTTACAGACAGAAAGAT 1628
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 944 TCTCTTAGTAGGAGAGAGAGCTGACCTGCTTCTTCTAGAGACAGACAGCTGAGATCAT 1003
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1629 GGCATTATACGCTACTCTTGGAAAAATGAGAAAAATGAGACTATGAAAAATTCATGGAAC 1688
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1004 GGCATTGACGCGCAGTGGAAAGTAAAGCCGGAACGAGAACTATGAAAAGTTTCATGAGAA 1063
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1689 AATGCGTAGCGCTTACTTTTGAATG 1715
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1064 AATGCGTAGCGCTTACTTCTGCTG 1090
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 14
AC158770/c      188707 bp DNA linear HTG 23-MAR-2005
LOCUS          Mus musculus chromosome 1 clone RP23-393018 map 1, WORKING DRAFT
DEFINITION     AC158770
SEQUENCE       AC158770.1 GI:61696497
VERSION        HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS       Mus musculus (house mouse)
SOURCE         Mus musculus
ORGANISM       Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
TITLE          Sciurognathi; Muroidae; Muridae; Murinae; Mus.
JOURNAL        1 (bases 1 to 188707)
REFERENCE      2 (bases 1 to 188707)
AUTHORS        Birren, B., Nussbaum, C. and Lander, E.
TITLE          Mus musculus chromosome 1, clone RP23-393018
JOURNAL        Unpublished
AUTHORS        2 (bases 1 to 188707)

```

Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B., Camarata, J., Chang, J., Choepel, T., Collymore, A., Cook, A., Cooke, P., Corum, B., D'Arrellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hages, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., MacDonald, P., Major, J., Manning, J., Matthews, C.,

TITLE
JOURNAL
COMMENT

McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Rachupka, A., Ramaamy, U., Raymond, C., Retta, R., Rice, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tsifaye, S., Theodore, J., Topham, K., Travers, M., Vaseillev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (23-MAR-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) <http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Broad Institute of MIT and Harvard

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@broad.mit.edu

Project Information

Center project name: 131651
Center clone name: 393_O_18

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731

Consensus quality: 183196 bases at least Q40
Consensus quality: 185063 bases at least Q30
Consensus quality: 186541 bases at least Q20

Insert size: 230000; agarose-fp
Insert size: 187207; sum-of-contigs

Quality coverage: 4.5 in Q20 bases; agarose-fp
Quality coverage: 5.5 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 16 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	52662: contig of 52662 bp in length
52663	52762: gap of unknown length
52763	52862: contig of 100 bp in length
52863	52962: gap of unknown length
52963	53916: contig of 954 bp in length
53917	54016: gap of unknown length
54017	56095: contig of 2079 bp in length
56096	56195: gap of unknown length
56196	57892: contig of 1697 bp in length
57893	57992: gap of unknown length
57993	61597: contig of 3605 bp in length
61598	61697: gap of unknown length
61698	66421: contig of 4724 bp in length
66422	66521: gap of unknown length
66522	70789: contig of 4268 bp in length
70790	70889: gap of unknown length
70890	74949: contig of 4060 bp in length
74950	75049: gap of unknown length
75050	81957: contig of 6908 bp in length
81958	82057: gap of unknown length
82058	89876: contig of 7821 bp in length
89879	89976: gap of unknown length
89979	103099: contig of 13121 bp in length
103100	103199: gap of unknown length
103200	114147: contig of 10948 bp in length
114148	114247: gap of unknown length
114248	130836: contig of 16591 bp in length
130839	130936: gap of unknown length
130939	162497: contig of 31559 bp in length

FEATURES	*	162498	162597	gap of unknown length
	*	162598	168707	contig of 26110 bp in length
source			Location/Qualifiers	
			1. 188707	
			/organism="Mus musculus"	
			/mol_type="genomic DNA"	
			/db_xref="taxon:10090"	
			/chromosome="1"	
			/map="1"	
			/clone="RP23-393018"	
			/clone_1fb="RPC1-23 Female Mouse BAC"	
misc_feature			1. 52662	
			/note="assembly_fragment	
gap			clone_end:SP6	
			vector_side:left"	
gap			52663..52762	
			/estimated_length=unknown	
misc_feature			52763..52862	
			/note="assembly_fragment"	
gap			52863..52962	
			/estimated_length=unknown	
misc_feature			52963..53916	
			/note="assembly_fragment"	
gap			53917..54016	
			/estimated_length=unknown	
misc_feature			54017..56095	
			/note="assembly_fragment"	
gap			56096..56195	
			/estimated_length=unknown	
misc_feature			56196..57892	
			/note="assembly_fragment"	
gap			57893..57992	
			/estimated_length=unknown	
misc_feature			57993..61597	
			/note="assembly_fragment"	
gap			61598..61697	
			/estimated_length=unknown	
misc_feature			61698..66421	
			/note="assembly_fragment"	
gap			66422..66521	
			/estimated_length=unknown	
misc_feature			66522..70789	
			/note="assembly_fragment"	
gap			70790..70889	
			/estimated_length=unknown	
misc_feature			70890..74949	
			/note="assembly_fragment"	
gap			74950..75049	
			/estimated_length=unknown	
misc_feature			75050..81957	
			/note="assembly_fragment"	
gap			81958..82057	
			/estimated_length=unknown	
misc_feature			82058..89878	
			/note="assembly_fragment"	
gap			89879..89978	
			/estimated_length=unknown	
misc_feature			89979..103093	
			/note="assembly_fragment"	
gap			103100..103199	
			/estimated_length=unknown	
misc_feature			103200..114147	
			/note="assembly_fragment"	
gap			114148..114247	
			/estimated_length=unknown	
misc_feature			114248..130838	
			/note="assembly_fragment"	
gap			130839..130938	
			/estimated_length=unknown	
misc_feature			130939..162497	
			/note="assembly_fragment"	
gap			162498..162597	
			/estimated_length=unknown	

Query Match	2.4%;	Score 56.6;	DB 12;	Length 188707;
Best Local Similarity	58.5%;	Pred. No. 9.7e-06;		
Matches 121;	Conservative	0;	Mismatches 79;	Indels 7; Gaps 1;
QY	1516	TGAGCTTTAGCCAGCACATCATCATGTAATATGCTTTCCTGATAGACCTGTTCTTAAT	1575	
DB	83324	TGAACCTTGCACCTTCCACATCACTACAGATGATGATTTGGTTCAGAGATTAAGAAATAGAAATTAAT	83465	
QY	1576	TCTCTTTGCAAAAGCTCTGCTACTTACTTACCAAGATC-----TGCCTACAGACAGAAAGAT	1628	
DB	83464	TCTCTCTAAGTGAACAGGACTGGAACCTCTCTCTTCTCTTAGAGACACACAGCTGAGATTCAT	83405	
QY	1629	GGCATTTTAACGGTACTTTGGAAAAATTAGAGAAAAATGAACTATGAAAAATTCATGGAGGC	1688	
DB	83404	GGCATTTTCGAAGGGAACGTGGAATAAGACCGGAACGGAATCTATGAAAAATTCATGGAGAA	83345	
QY	1689	AATGGGTAAAGCCTTACTTTTTCATATG	1715	
DB	83344	AATGGGTAAAGCACTGGCTTCTCTCTG	83318	

[illegible]

TITLE
JOURNAL
REFERENCE
AUTHORS

Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karataas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramaamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (22-Oct-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 240767)
Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karataas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramaamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted (10-Nov-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA
5 (bases 1 to 240767)
Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karataas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramaamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

FEATURES
source
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Broad Institute of MIT and Harvard
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: U30890
Center clone name: 70_L_3

Location/Qualifiers
1..240767
/organism="Mus musculus"
/mol_type="Genomic DNA"
/db_xref="taxon:10090"
/chromosome="3"
/map="3"
/clone="RP24-70L3"
/clone_id="RPCI-24 Male Mouse BAC"
110..298
/rpt_family="L2"
complement(11767..2017)
/rpt_family="B4A"
2076..2104
/rpt_family="AT_rich"
2157..2205
/rpt_family="AT_rich"
2305..2398
/rpt_family="Lx4"
complement(3241..3329)
/rpt_family="L1MEC"
3371..3457
/rpt_family="PB1D10"
3699..3738
/rpt_family="T-rich"
3825..3942
/rpt_family="B1_MM"
complement(4160..4441)
/rpt_family="L1MEC"
complement(4648..4730)
/rpt_family="MIR"
complement(4768..4850)
/rpt_family="MIR"
5595..5637
/rpt_family="T-rich"
complement(6332..6443)
/rpt_family="L2"
complement(6768..7027)
/rpt_family="B4A"
7531..7581
/rpt_family="AT_rich"
complement(7649..7898)
/rpt_family="Lx8"
8014..8135
/rpt_family="B4"
8197..8370
/rpt_family="B2_MM2"
8725..8753
/rpt_family="(T)n"
10512..10618
/rpt_family="PB1D7"
10654..10845
/rpt_family="RME17C"
11363..11475
/rpt_family="PB1"
complement(11517..11546)
/rpt_family="B3"
complement(11547..11695)
/rpt_family="B1_MM"
complement(11696..11877)
/rpt_family="B3"
complement(11894..11987)
/rpt_family="B1F"

```
repeat_region complement(13781..13925)
/rpt_family="B1F"
repeat_region 13983..14157
/rpt_family="(TG)n"
repeat_region 14188..14520
/rpt_family="L1ME1"
unsure 14247
/note="single clone coverage"
repeat_region 14615..14642
/rpt_family="(A)n"
```

Query Match 2.4%; Score 56.6; DB 6; Length 240767;

Best Local Similarity 58.5%; Pred. No. 9.8e-06; Mismatches 79; Indels 7; Gaps 1;

Matches 121; Conservative 0;

```
QY 1516 TGAGCTTTAGCCAGCCATCATGTAATTGCTTCTCTGATTAAGCTGTTCATTAAT 1575
DB 74906 TGAACCTTCGAACTTCACATCAGATGATGATTGTTGGAAGATTAAGAAATAGATTAAT 74847
QY 1576 TCTCTTGCAGAGCTCTGCTACTTACCGAAGTC-----TGCCTACAGACAGAAAGAT 1628
DB 74846 TCTCTTAGTGAGACAGACTGACCTCTGCTTCTCTAGAGACACACAGCTGATCAT 74787
QY 1629 GGCATTTAAGGATCTTGAAGAAATAGAGAAATAGAACTATGAAATTCATGGAAGC 1688
DB 74786 GGCATTCGACGGCAGTGAGAAAGTAGACCGGACGAGAACTATGAAAGTTCAATGAGAA 74727
QY 1689 AATGGTAAAGCCTTACTTTTGAATG 1715
DB 74726 AATGGTAAAGCACTGGCTTCTGCTG 74700
```

Search completed: May 28, 2006, 10:39:22
Job time : 13882.8 secs

CC gut-specific gene expression control region operably linked to a nucleic
CC insert encoding a polypeptide, an expression vector that integrates
CC into a host cell (and comprising the isolated avian gut-specific gene
CC expression control region), expressing a heterologous polypeptide in a
CC host cell (by transfecting a eukaryotic cell with the recombinant DNA
CC molecule, and culturing the transfected cell in a medium suitable for
CC expression of a heterologous polypeptide under the control of an avian
CC intestinal fatty acid binding protein (IFABP) or cp15 gene expression
CC control region encoded by the recombinant DNA molecule), a eukaryotic
CC cell transformed with the expression vector (or its progeny, which
CC expresses a heterologous polypeptide) and a transgenic avian having a
CC heterologous polynucleotide sequence comprising the nucleic acid insert.
CC The nucleic acids are useful for regulating heterologous nucleic acid in
CC transgenic avians, as probes in nucleic acid hybridisation assays for
CC detecting the IFABP gene expression control region, and for generating
CC transgenic birds. The present sequence is the chicken intestinal fatty
CC acid binding protein, IFABP, gene, 5' region.

SQ Sequence 2381 BP; 775 A; 428 C; 469 G; 709 T; 0 U; 0 Other;

Query Match 100.0%; Score 2381; DB 11; Length 2381;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTTCCTGCGCAGAAAAGGCTGTGGGGTCTTGTTCCTCAGACGCTTAAGCAAAATC 60
DB 1 AGCTTCCTGCGCAGAAAAGGCTGTGGGGTCTTGTTCCTCAGACGCTTAAGCAAAATC 60
QY 61 CCCAAGTTCAAAACGTCGGCTGTAAAGAGAGATGGCTCACTTCAAAATGAATGATG 120
DB 61 CCCAAGTTCAAAACGTCGGCTGTAAAGAGAGATGGCTCACTTCAAAATGAATGATG 120
QY 121 AAATATCATTAACGAGCTCTGTTGGCAGATCAGAGATTAACCTTGTGGGAGCAAAAT 180
DB 121 AAATATCATTAACGAGCTCTGTTGGCAGATCAGAGATTAACCTTGTGGGAGCAAAAT 180
QY 181 CTTAAAGTGATAGGTAGAACGAGGTGTGCAACTTAATGACTAGATTGACACA 240
DB 181 CTTAAAGTGATAGGTAGAACGAGGTGTGCAACTTAATGACTAGATTGACACA 240
QY 241 TTTGATCTTCTAGAGACAAAAGGCTCTGGAACCAATTAATCTGGTGACAGTCACT 300
DB 241 TTTGATCTTCTAGAGACAAAAGGCTCTGGAACCAATTAATCTGGTGACAGTCACT 300
QY 301 AGCAGCCCTGTTGGGTGCACTACAGCACTTTGTTGCAACAATTAATCTGAAGTTG 360
DB 301 AGCAGCCCTGTTGGGTGCACTACAGCACTTTGTTGCAACAATTAATCTGAAGTTG 360
QY 361 TTTTCTTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 420
DB 361 TTTTCTTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 420
QY 421 TACTTCCCTGCGACCTCTTACAGCAATTAAGCAATTAAGCAATTAAGCAATTAAGCA 480
DB 421 TACTTCCCTGCGACCTCTTACAGCAATTAAGCAATTAAGCAATTAAGCAATTAAGCA 480
QY 481 ATGGAAGCCACACTATTTTGGTATCAGCAGCAAAAGTTCTAAATTAATTAATTAATTA 540
DB 481 ATGGAAGCCACACTATTTTGGTATCAGCAGCAAAAGTTCTAAATTAATTAATTAATTA 540
QY 541 GTAAACCTTTTGTGCGAGCTCTGGAAGAAAAGAAATTAATTAATTAATTAATTAATTA 600
DB 541 GTAAACCTTTTGTGCGAGCTCTGGAAGAAAAGAAATTAATTAATTAATTAATTAATTA 600
QY 601 AAATTCATTAAGTAAAGCCCAACCTGTTATCTGTGATTAAGCAATTAATTAATTAATTA 660
DB 601 AAATTCATTAAGTAAAGCCCAACCTGTTATCTGTGATTAAGCAATTAATTAATTAATTA 660
QY 661 TCAGAGATTACATTTTGAAGTAAATGACAGAGATTTGGTCCGTCATTAAGAAACGA 720
DB 661 TCAGAGATTACATTTTGAAGTAAATGACAGAGATTTGGTCCGTCATTAAGAAACGA 720
QY 721 CTAACTATATCTGAGTTTATGACAAAGATTTAGCAACAGCAAAATTTGCTCAGTTTC 780

DB 721 CTAACTATATCTGAGTTTATGACAAAGATTTAGCAACAGCAAAATTTGCTCAGTTTC 780
QY 781 AAGTAGACATATCTGAGTTTATGACAAAGATTTAGCAACAGCAAAATTTGCTCAGTTTC 840
DB 781 AAGTAGACATATCTGAGTTTATGACAAAGATTTAGCAACAGCAAAATTTGCTCAGTTTC 840
QY 841 ATCTTCAGCAATTTAAACCTGGGATCTATGAAATCAAAACAGTTGGGTAATTAATCACT 900
DB 841 ATCTTCAGCAATTTAAACCTGGGATCTATGAAATCAAAACAGTTGGGTAATTAATCACT 900
QY 901 TAGGAGACATCACTACTGTAGAAATGACAGAAACAGAGATTTCACTGAATGGCTAT 960
DB 901 TAGGAGACATCACTACTGTAGAAATGACAGAAACAGAGATTTCACTGAATGGCTAT 960
QY 961 AATATAGAGATAGTAGAAGGTGCTGAAATTTAGACTAATTAATTAAGATAGAGACA 1020
DB 961 AATATAGAGATAGTAGAAGGTGCTGAAATTTAGACTAATTAATTAAGATAGAGACA 1020
QY 1021 GGAATGAGAAATATCATGCAATTTCTGTAGCTCAGCACTAGACTGAAAGTTTCTGAAA 1080
DB 1021 GGAATGAGAAATATCATGCAATTTCTGTAGCTCAGCACTAGACTGAAAGTTTCTGAAA 1080
QY 1081 CTGAACCGAGTTTCCCAAACTACCTGTGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 CTGAACCGAGTTTCCCAAACTACCTGTGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 TATGTGAGTAGAATAGATTCTCACCAATTAAGATGACAAAGAGATTTGTGTTTT 1200
DB 1141 TATGTGAGTAGAATAGATTCTCACCAATTAAGATGACAAAGAGATTTGTGTTTT 1200
QY 1201 ATCTGTGGGTAAATAGCTTTCTCAGTGTATTAACACCTCCACAGATTAAGATC 1260
DB 1201 ATCTGTGGGTAAATAGCTTTCTCAGTGTATTAACACCTCCACAGATTAAGATC 1260
QY 1261 CTATGCAACAAAGAAATGTCATATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1320
DB 1261 CTATGCAACAAAGAAATGTCATATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1320
QY 1321 GCGGTTTTTTCTTCAACTCAATTAAGATGACAAAGATGATGATGATGATGATGAT 1380
DB 1321 GCGGTTTTTTCTTCAACTCAATTAAGATGACAAAGATGATGATGATGATGATGAT 1380
QY 1381 TAAAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440
DB 1381 TAAAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440
QY 1441 CAAACAGTGTCTTCAAGTGGAAACAAAGTTAACTAAAGTCAACCCCTCTTGAACA 1500
DB 1441 CAAACAGTGTCTTCAAGTGGAAACAAAGTTAACTAAAGTCAACCCCTCTTGAACA 1500
QY 1501 GATCAATGCGACAGTGGAGCTTTAGCCAGCACTCATATTAATTAATTAATTAATTAAT 1560
DB 1501 GATCAATGCGACAGTGGAGCTTTAGCCAGCACTCATATTAATTAATTAATTAATTAAT 1560
QY 1561 AGCTGTTCATTAATTTCTTTTGAAGAGCTCTGCTACTTCAAGAGTCTGCTTACAGAC 1620
DB 1561 AGCTGTTCATTAATTTCTTTTGAAGAGCTCTGCTACTTCAAGAGTCTGCTTACAGAC 1620
QY 1621 AGAAGATGCAATTTAAGCTTCTTGAAGAAATAGAGAAATAGAGAACTATGAAAAATTC 1680
DB 1621 AGAAGATGCAATTTAAGCTTCTTGAAGAAATAGAGAAATAGAGAACTATGAAAAATTC 1680
QY 1681 ATGGAACAAATGGGTAAAGCTTACTTTTGAATGCTCTTCAAAACAGAGATCACTAC 1740
DB 1681 ATGGAACAAATGGGTAAAGCTTACTTTTGAATGCTCTTCAAAACAGAGATCACTAC 1740
QY 1741 GCGGATTAACAACTTAAGCTTCTTCAATGAATCAATCTGCTTAACCTTCTTCTTCTT 1800
DB 1741 GCGGATTAACAACTTAAGCTTCTTCAATGAATCAATCTGCTTAACCTTCTTCTTCTT 1800
QY 1801 CTGCTATTTTGGCTTGCACATGCTGCTGACTTATTTGAAAAAGCTTATTAAGGGGA 1860

Db	1801	CTGCAATTTTGGCCTTGGACATTTGCCCTGCACCTTATTTTGGAAAAGACTCTATAGAGGGGA	1866
Qy	1861	ATACAGAGAAAGAAAACATTTCTGATTTTATTTTGCAATTTGGCAATCTTATGCAATTAGCT	1920
Db	1861	ATACAGAGAAAGAAAACATTTCTGATTTTATTTTGCAATTTGGCAATCTTATGCAATTAGCT	1920
Qy	1921	AATTCAGTAGAGGCAATTCACAGAGAAATTTAAATAGAAATATATGTAGGAATTTATTT	1980
Db	1921	AATTCAGTAGAGGCAATTCACAGAGAAATTTAAATAGAAATATATGTAGGAATTTATTT	1980
Qy	1981	TTGATTAAGACTGTTTGAAAATAATACAGAGAGGAAATTTGCTGTCTCCAGTTTTCAGA	2040
Db	1981	TTGATTAAGACTGTTTGAAAATAATACAGAGAGGAAATTTGCTGTCTCCAGTTTTCAGA	2040
Qy	2041	ACACACATGATTTTGAGTCATTTTAAACATGCTAGTCTTACTTTAAGCTTTGACAACTGC	2100
Db	2041	ACACACATGATTTTGAGTCATTTTAAACATGCTAGTCTTACTTTAAGCTTTGACAACTGC	2100
Qy	2101	CTGTATATATGAGATGTAAACATTAACATCTCTAGTTGGATAGATAGTTTGTATTAACAGGCTGA	2160
Db	2101	CTGTATATATGAGATGTAAACATTAACATCTCTAGTTGGATAGATAGTTTGTATTAACAGGCTGA	2160
Qy	2161	ACACTGCTCAGTGAAGAGTGGAGAGAGTAAGACTCTGAGTCAGAAATTCGGCTTAAGC	2220
Db	2161	ACACTGCTCAGTGAAGAGTGGAGAGAGTAAGACTCTGAGTCAGAAATTCGGCTTAAGC	2220
Qy	2221	TCCCTGCACTACAGAAAAAGTCACATATATAAATGCAACATGATGTTCTATTTTGT	2280
Db	2221	TCCCTGCACTACAGAAAAAGTCACATATATAAATGCAACATGATGTTCTATTTTGT	2280
Qy	2281	CTCTGCTTGATGTTTATTTGATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	2340
Db	2281	CTCTGCTTGATGTTTATTTGATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	2340
Qy	2341	GAAAGTTAGAGCCCAAGCATATCTGAAGCTCAGCTATTCAG	2381
Db	2341	GAAAGTTAGAGCCCAAGCATATCTGAAGCTCAGCTATTCAG	2381
RESULT 2			
ADL90128			
ID	ADL90128	standard; DNA; 336 BP.	
AC	ADL90128;		
XX			
DT	20-MAY-2004	(first entry)	
XX			
DE	Chicken intestinal fatty acid binding protein, <i>IFABP</i> , gene, promoter.		
XX			
KW	Chicken; <i>de</i> ; intestinal fatty acid binding protein; <i>IFABP</i> ;		
KW	gut specific promoter; transgenic; promoter.		
XX			
OS	Gallus gallus.		
XX			
PN	US2003177516-A1.		
XX			
PD	18-SEP-2003.		
XX			
PE	14-MAR-2002; 2002US-00099663.		
XX			
PR	14-MAR-2002; 2002US-00099663.		
XX			
PA	(HORS// HORSEMAN N D.		
XX	(PRATT// PRATT S L.		
XX			
PI	Horseman ND, Pratt SL;		
XX			
DR	WPI; 2003-898653/82.		
PT	New nucleic acid molecule comprising an isolated avian gut-specific gene		
PT	expression control region, useful for regulating heterologous nucleic		
XX	acids in transgenic avians, and for generating transgenic birds.		

FT misc_feature 3890 /tag= 1 /note= "Represented as * in the specification"
 FT misc_feature 3912 /tag= j /note= "Represented as * in the specification"
 FT misc_feature 3914 /tag= k /note= "Represented as * in the specification"
 FT misc_feature 3938 /tag= l /note= "Represented as * in the specification"
 FT misc_feature 3939 /tag= m /note= "Represented as * in the specification"
 FT misc_feature 3941 /tag= n /note= "Represented as * in the specification"
 FT misc_feature 3943 /tag= o /note= "Represented as * in the specification"
 FT misc_feature 4361 /tag= p /note= "Represented as * in the specification"
 FT misc_feature 4361 /tag= q /note= "Represented as * in the specification"
 WO200133977-A1.
 17-MAY-2001.
 06-NOV-2000; 2000WO-AU001362.
 05-NOV-1999; 99AU-00003875.
 (META-) METABOLIC PHARM LTD.
 Belyea CI, Ng FM, Vaughan P;
 WPI; 2001-328876/34.
 New organisms containing nucleic acid encoding a growth hormone fragment which modulates lipid metabolism are useful to produce dietary aids for obesity and in the meat production industry.
 Disclosure; Page 48-50; 54pp; English.
 The invention relates to novel transgenic organisms useful in the production of functional food and drink products for the treatment or prevention of obesity via the regulation of lipid metabolism. The organisms comprise a polynucleotide encoding a growth hormone fragment capable of stimulating the activity of hormone-sensitive lipase (the key enzyme in lipolysis) and inhibiting acetyl CoA carboxylase (the key enzyme in lipogenesis). The growth hormone fragment preferably contains at least the disulphide-bonded loop of a mammalian growth hormone (but is not the full-length growth hormone) and is optionally linked to an epitope tag or heterologous fusion protein partner. The transgenic organism may be a microorganism used to produce a fermented product (e.g., yeast), or an edible plant or animal or cell thereof. Food or drink made using methods of the invention are used to modify fat/lean ratio, lipid metabolism or food use in a mammal. In particular, the food or drink products may be used to treat or prevent obesity, particularly in humans, and may also be used to improve the fat/lean ration of livestock raised for meat production. In the exemplification of the invention, the human growth hormone (hGH) fragment analogue AOD9604 was expressed in yeast, optionally fused to the FLAG epitope (ABR73625). The present sequence is described as a DNA sequence from yeast in the CC sequence listing, but is not further referred to in the specification
 SX Sequence 4590 BP; 661 A; 384 C; 127 G; 522 T; 0 U; 2896 Other;
 Query Match 2.1%; Score 50.6; DB 5; Length 4590;
 Best Local Similarity 11.8%; Pred. No. 0.19; Mismatches 423; Indels 3; Gaps 3; Matches 113; Conservative 416; Mismatches 423; Indels 3; Gaps 3;

QY 153 CAGAGATTAACCTGCTGGGCAAAATCTTAAAGTATAGAGAAACAGAGCTGTT 212
 DB 3544 CCRDNGTCAMBSBANTATYASHRSDSCBDDTBRNCTTHAMNSANDRAMGSDATDHS 3603
 QY 213 GCAACTAAATGCACTAGATTCTTACCACATTTGATCTCTAGAGACAAAGGCTGGAA 272
 DB 3604 NNTDADTADYAGSTNNGRMATTHDAAVNNTRNTHANDIARCTBGRNMAATNMBNKN 3663
 QY 273 AACAAATTAATTGTCGACAGTCAGTAGACGCTGTTTGGGTCACATACCAACT 332
 DB 3664 TRACANTSNAMACSNMGRMATDARSARHANDVWNTCTDAND -RAATNVRSTYATH 3722
 QY 333 TTGTTGCAACAATTAACAATCTAAGTGTCTTTCTCTCTCTTCTTAATCTCTGACA 392
 DB 3723 RTYADRSARHANDNSTRADVWNTDHYTHATNTATNTYSGRTHACKMSNSTTGA 3782
 QY 393 GTCTTAAGGTGAAGAGTACTGATTAAGTACTTCCCTCGATCTCTTACCCAGATTAG 452
 DB 3783 TGTADTADTADVARAATRTNYSASTRAAAGAGANMNRGMATBMSHMNDYSKCTMD 3842
 QY 453 CATTGATTTCAAAATGAACCTGAGTGAAGAACCAACATATTTGATGACACGAC 512
 DB 3843 MMYTKCAKCDMNTGHCNHNBRBDHDDDBVMBRVHDCDVBNBNVDDHDDH 3902
 QY 513 -AAAGTTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 571
 DB 3903 DHDNDMCHNNBVBVHDDHDDVYVMNRGHNHNNHNNHTCDHDDHNTCHABDG 3962
 QY 572 AAGAAATTAATGAATTAACCAAGTACTGATTAATTAATTAATTAATTAATTAATTAAT 631
 DB 3963 MVVVVVVVVVVVVVVCHKHDDHDKHMKCMHCCCHKMDHAAANMAAMHNNHNDG 4022
 QY 632 TGTGATTAAGCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 691
 DB 4023 CKHADATRARARANTTADRTTRAGNBRNMBSCDMCKMMKRYKDBDWRYTT 4082
 QY 692 GCAGATTTGGTCCGTCATAGAACAGACTAATTAATCTGATTTAGTACAGACG 751
 DB 4083 GBDKQIMNDYSMSHMNDYSKCMNDYSTMAATRCVRSVAVASDMNDYSTMAATRCV 4142
 QY 752 ATTTAGCAACCAATTTGCTGATTTGCTGATTTGCTGATTTGCTGATTTGCTGATTTG 811
 DB 4143 YSAVASDMNDYSTMAATRCVRSVAVASD-WNDYSTMAATRCVRSVAVASDMNDYSTMA 4201
 QY 812 TGACCAAGTNGTCCATTTCTGATTTCTTCAATTTCAATTTCAATTTCAATTTCAATTT 871
 DB 4202 RCVYSHVAVASDMNDYSTMAATRCVRSVAVASDRKSKMWSKMSMGTMNMBANSTYBA 4261
 QY 872 GAATTCAAACGCTGGGTAAATTTCACTTACGACACATCACTCACTGATGAAATGAC 931
 DB 4262 RACRBNHNBBSDBMSSTDMRDSCSTMAATSCDTMNDYSNDYSHHSNMNDYNDNDSC 4321
 QY 932 AGAAACAGACATTCAGTAAATGGCTATTAATTAATTAATTAATTAATTAATTAATTAAT 991
 DB 4322 WNDYSDMCRSTWRBRBDGTHNAKCHTTDGBNMBDDRTNTRSHARSRTABAC 4381
 QY 992 TTTAGTACCTAATTAAGAGTGAAGACAGAAATGAGAAATATGATGCAATTTCTGAG 1051
 DB 4382 ATTSADTTADNSTDMDHDDSTWRDCCNTCMNTSKCNTMTDNTMAASADGSTPMARY 4441
 QY 1052 CTCAGCACTAGACTGGAAGTTCTGAAATGAAACGAGTTTCCCAATCACTG 1106
 DB 4442 NRYATNNNSRTATTCANAGARRVYDCNMTSMARVYNNMATRDSCTMTCDACMS 4496
 RESULT 5
 ID ABR83946 standard; cDNA; 5204 BP.
 XX ABR83946;
 XX AC ABR83946;
 XX DT 14-AUG-2002 (first entry)
 XX

DE Human cDNA differentially expressed in granulocytic cells #517.
XX
XX Human; se; granulocytic cell; DNA chip; bacterial infection;
KM viral infection; parasitic infection; protozoal infection;
KM fungal infection; sterile inflammatory disease; psoriasis;
KM rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KM cardiac reperfusion injury; renal reperfusion injury; ARDS;
KM adult respiratory distress syndrome; inflammatory bowel disease;
KM Crohn's disease; ulcerative colitis; periodontal disease;
KM granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-Apr-2002.
XX
PF 03-OCT-2001; 2001MO-US030821.
XX
PR 03-OCT-2000; 2000US-0237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weisman SM, Yamaga S, Vockley J;
XX WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression of
PT genes associated with granulocyte activation, which serves as diagnostic
PT markers that is useful for monitoring disease states and drug toxicity.
XX
PS Claim 1, SEQ ID NO 517; 114pp; English.
XX
XX The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing the
CC expression level to an expression level in an unactivated GC, where
CC differential expression of Gs is indicative of GCA. Also included are
CC modulating (M2) GA by contacting GC with an agent that alters the
CC expression of at least one gene in Gs; (2) screening (M3) for an agent
CC capable of modulating GCA or an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease using the gene expression
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from Gs, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC for screening an agent capable of modulating GCA preferably in an
CC inflammation in a tissue; M4 is useful for detecting an inflammation
CC (especially chronic) in a tissue, an allergic response in a subject,
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, periodontal disease, also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 5204 BP; 1770 A; 867 C; 836 G; 1731 T; 0 U; 0 Other;

Query Match 2.1%; Score 49.8; DB 6; Length 5204;
Best Local Similarity 59.2%; Pred. No. 0.31;

Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;
QY 1516 TGAGCTTAGCAGCCAGCATCATATGTAATGCTTCTGTGATGAAGCTGTCATTAAT 1575
DB 971 TGAACCTTTAAGCTTCCACATCACAGATATGAAAGTTGGTTCAGATGAATAATTAATTAAT 1030
QY 1576 TCTC-----TTTGCAAGCTCTGCTACTTACCAAGAGCTGCTACAGACAGAAAGA 1627
DB 1031 TCTGCCCCAAGACAGACCGTGAATCTCTAGCTCTGAGGCTGACT-CAACTGAATATCA 1089
QY 1628 TGGCATTTTAACGCTACTTGGAAAAATAGAAAAATGAGAACTATGAAAAATTCATGAGAG 1687
DB 1090 TGGGCTTTGACACACACTGGAGAGTAGACCGGAGTGAATAACTATGACAGAGTTTCATGAGAA 1149
QY 1688 CAATGGTAAAGCCTTACTTTTGAATGCT 1718
DB 1150 AATGGGTAAAGACTTATTTCTTGTGGCT 1180

RESULT 6
ADJ74588
ID ADJ74588 standard; DNA; 5204 BP.

XX AC ADJ74588;

XX DT 06-MAY-2004 (first entry)

XX DE Human intestinal fatty acid binding protein gene SEQ ID NO:12.

XX KM db; gene; human; intestinal fatty acid binding protein; restenosis;
KM coronary angioplasty; balloon coronary angioplasty;
KM stent coronary angioplasty.

XX OS Homo sapiens.

XX PN WO2004015104-A1.

XX PD 19-FEB-2004.

XX PF 20-MAR-2003; 2003WO-JP003478.

XX PR 09-AUG-2002; 2002JP-00233041.

XX PA (NAGO-) NAGOYA IND SCI RES INST.
XX (GIFU-) GIFU INT INST BIOTECHNOLOGY.

XX PI Yamada Y, Yokota M;

XX DR WPI; 2004-180672/17.

XX PT Analysis of specific gene polymorphisms in clinical nucleic acid sample
PT for prediction of risk of restenosis after balloon or stent coronary
PT angioplasty.

XX PS Disclosure; SEQ ID NO 12; 164pp; Japanese.

XX CC The invention relates to a novel method for predicting the risk of
CC restenosis after coronary angioplasty comprising analyzing specific gene
CC polymorphisms in a clinical nucleic acid sample. The method is useful for
CC the diagnosis of the genetic risk of restenosis following balloon or
CC stent coronary angioplasty. The method has high accuracy and high
CC estimation ratio. The present sequence is used in the exemplification of
CC the invention.

XX SQ Sequence 5204 BP; 1770 A; 867 C; 836 G; 1731 T; 0 U; 0 Other;

Query Match 2.1%; Score 49.8; DB 12; Length 5204;
Best Local Similarity 59.2%; Pred. No. 0.31;
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

QY 1516 TGAGCTTAGCAGCCAGCATCATATGTAATGCTTCTGTGATGAAGCTGTCATTAAT 1575
DB 971 TGAACCTTTAAGCTTCCACATCACAGATATGAAAGTTGGTTCAGATGAATAATTAATTAAT 1030

Qy 1576 TCTC-----TTTGAAAGCTCTGCTACTTACCAAGTCTGCTTACAGACAGAAAGA 1627
 Db 1031 TCTGCCCCAGAGACAGACTGTAATCTTAGCTGCTAGAGGCTGACT-CAACTGAAATCA 1089
 Qy 1628 TGGCATTTAACGCTACTTGGAAATAGAGAAATAGAACTATGAAATTCATGGAAG 1687
 Db 1090 TGGCGTTTGAACAGCACTTGGAAAGGTAGACCGAGTGAATACTATGACAACTCATGAAA 1149
 Qy 1688 CAATGGGTAGCCTTACTTTTGTGAATGCT 1718
 Db 1150 AAATGGTAAAGACTTATTTCTTTGGCT 1180

RESULT 7
 ID ADO78215 standard; DNA; 5204 BP.
 AC ADO78215;
 DT 29-JUL-2004 (first entry)
 DE Human fatty acid binding protein 2 Alu repeat region DNA with SNP Seq 9.
 KW human; de; single nucleotide polymorphism; SNP; hereditary risk;
 KW coronary spasm; tumor necrosis factor alpha;
 KW NADH/NADPH oxidase p22 phox gene; apolipoprotein B; apolipoprotein C-III;
 KW stromelysin 1; interleukin 6;
 KW endothelium type nitrogen monoxide synthetase;
 KW fatty acid binding protein 2; cardiac sudden death;
 KW variant angina pectoris; myocardial infarction.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT variation rebase(2445.A)
 FT /*tag= a
 FT /standard_name= "Single nucleotide polymorphism"
 PN JP2004113093-A.
 XX 15-APR-2004.
 XX 25-SEP-2002; 2002JP-00280031.
 XX 25-SEP-2002; 2002JP-00280031.
 XX (NAGO-) ZH NAGOYA SANGYO KAGAKU KENKYUSHO.
 PA (GIFU-) ZH GIFUKEN KOKUSAI BIO KENYUO.
 DR WPI; 2004-360127/34.
 PT Detecting genotype in nucleic acid sample useful for diagnosing risk of
 PT coronary spasm, by analyzing polymorphisms in tumor necrosis factor alpha
 PT gene, NADH/NADPH oxidase p22 phox gene and apolipoprotein B gene, in
 PT nucleic acid sample.
 PS Disclosure; SEQ ID NO 9; 59pp; Japanese.
 CC This invention relates to a novel method for detecting a hereditary risk
 CC for a coronary spasm. Specifically, it refers to the identification of
 CC two or more polymorphisms and determining the genotype of the nucleic
 CC acid sample in order to calculate the risk of the patient to a coronary
 CC spasm. The present invention describes a risk analysis that comprises
 CC identifying two or more polymorphisms occurring at any position taken
 CC from -863 of the tumor necrosis factor alpha gene, 242 of the NADH/NADPH
 CC oxidase p22 phox gene, -219 of the apolipoprotein B gene, -482 of the
 CC apolipoprotein C-III gene, -1171 of the stromelysin 1 gene, 4070 of the
 CC apolipoprotein B gene, -634 of the interleukin 6 gene, -786 of the
 CC endothelium type nitrogen monoxide synthetase gene or position 2445 of
 CC the fatty acid binding protein 2 gene. Accordingly, the method enables
 CC prevention of cardiac sudden death resulting from variant angina
 CC pectoris, coronary spasm and myocardial infarction. This polymorphisms

CC is the human fatty acid binding protein 2 gene Alu repeat region
 CC containing a G2445A SNP in the sequence given in the invention. NOTE:
 CC This sequence is referred to as Seqid 9 in the sequence listing, but
 CC Seqid 8 also refers to a primer given in the disclosure of the
 CC specification.
 SQ Sequence 5204 BP; 1770 A; 867 C; 836 G; 1731 T; 0 U; 0 Other;
 Query Match 2.1%; Score 49.8; DB 12; Length 5204;
 Best Local Similarity 59.2%; Pred. No. 0.31;
 Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

Qy 1516 TGAGCTTTAGCCAGCCATCATCATGTAATGCTTCTGATAGCCTTCTAAT 1575
 Db 971 TGAACCTTAAGCTTCCACATCATCAGTATGAAGTTGTTCAAGATGAATAATTAAT 1030
 Qy 1576 TCTC-----TTTGAAAGCTCTGCTACTTACCAAGTCTGCTTACAGACAGAAAGA 1627
 Db 1031 TCTGCCCCAGAGACAGACTGTAATCTTAGCTGCTAGAGGCTGACT-CAACTGAAATCA 1089
 Qy 1628 TGGCATTTAACGCTACTTGGAAATAGAGAAATAGAACTATGAAATTCATGGAAG 1687
 Db 1090 TGGCGTTTGAACAGCACTTGGAAAGGTAGACCGAGTGAATACTATGACAACTCATGAAA 1149
 Qy 1688 CAATGGGTAGCCTTACTTTTGTGAATGCT 1718
 Db 1150 AAATGGTAAAGACTTATTTCTTTGGCT 1180

RESULT 8
 ID AED14682 standard; DNA; 426 BP.
 AC AED14682;
 DT 01-DEC-2005 (first entry)
 DE Mutant rat intestinal fatty acid binding protein DNA.
 KW high throughput screening; diagnosis; fluorescence; drug metabolism;
 KW cerebrovascular ischemia; cerebroprotective; vasotropic; heart disease;
 KW cardiac; cardiovascular disease; cancer; cytotoxic; neoplasm;
 KW intestinal fatty acid binding protein; I-FABP; mutant; de; gene.
 OS Rattus norvegicus.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT CDS 1..423
 FT /*tag= a
 FT /product= "intestinal fatty acid binding protein"
 PN WO2005093103-A2.
 XX 06-OCT-2005.
 XX 21-MAR-2005; 2005WO-US009278.
 XX 22-MAR-2004; 2004US-0555224P.
 XX (FPA-) FPA SCI LLP.
 PA Kleinfeld AM, Huber AH, Kampf JP, Kwan T, Zhu B;
 DR WPI; 2005-713873/73.
 DR P-PSDB; AED14683.
 PT High throughput generation and screening of probes to determine
 PT concentration of unbound bilirubin in, e.g. blood, comprises generating
 PT polymers, expressing proteins, purifying, associating with
 PT fluorophores, retrieving and screening.
 PS Disclosure; SEQ ID NO 3; 82pp; English.

XX The new invention relates to high throughput screening methods to provide
 CC specific probes that measure levels of unbound analytes, including
 CC unbound free fatty acids. Specifically claimed is a method of high
 CC throughput generation and screening of probes (M1) by generating
 CC polynucleotides encoding a protein library, expressing and purifying the
 CC proteins by binding to a solid matrix, associating the matrix bound
 CC proteins with fluorophores, retrieving the probes from the solid matrix,
 CC and screening the probes in a fluorometer. Also claimed are a method of
 CC determining the concentration of unbound bilirubin in body fluid of a
 CC mammal by withdrawing a body fluid, contacting the fluid with a probe,
 CC and determining the level of unbound bilirubin; a fluorescently labeled
 CC protein based peptide sequence (rat intestinal FABP) which comprises an
 CC additional mutation; a method of measuring an amount of an unbound drug
 CC in a mammal by withdrawing a body fluid, measuring the binding of the
 CC drug or its metabolite to a probe thus determining the amount of the
 CC unbound drug; a polynucleotide encoding a functional engineered protein;
 CC an ILBP capable of binding an unbound metabolite. In (M1) the template
 CC template protein is a non-enzymatic and the molecule is an unbound metabolite. The
 CC template protein is an intracellular Lipid Binding Protein (ILBP). The
 CC used for generating and screening of probes used for determining the
 CC concentration of unbound bilirubin in body fluids, e.g. blood, blood
 CC plasma, blood serum, and urine. The unbound bilirubin is used to
 CC determine the individual's relative risk for disease such as stroke,
 CC cardiac disease, and cancer. The present sequence is mutant rat
 CC intestinal fatty acid binding protein, DNA. The rat FABP is post
 CC translationally modified to remove the n-terminal methionine and
 CC acetylate the n-terminal residue alanine.

Sequence 426 BP; 145 A; 76 C; 110 G; 95 T; 0 U; 0 Other;

Query Match 2.1%; Score 49.2; DB 14; Length 426;
 Best Local Similarity 81.4%; Pred. No. 0.19;
 Matches 57; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1627 ATGCATTACGCTACTTGGAAATAGAGAAAATGAACTTGAAAATTCATGCA 1666
 DB 1 ATGCATTATGTCACCTTGGAAATAGAGAAAATGAACTTGAAAATTCATGCA 60

QY 1687 GCATGGGTA 1696
 DB 61 AAAATGGCA 70

RESULT 9
 ADA71938
 ID ADA71938 standard; DNA; 2000 BP.

AC ADA71938;

XX 20-NOV-2003 (first entry)

DE Rice gene, SEQ ID 5263.

XX Plant; bacterial infection; fungal infection; viral infection; rice;
 KM gene; ds.

XX Oryza sativa.

XX MO2003000896-A1.

PN 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

PR (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
 PI Katsagiri F, Qian S, Tao Y, Whitlam S, Xie Z, Zhu T, Zou G;
 XX

DR WPI; 2003-175290/17.
 XX Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX Claim 27; SEQ ID NO 5263; 899bp; English.
 XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant. In a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 2.1%; Score 49.2; DB 8; Length 2000;
 Best Local Similarity 10.8%; Pred. No. 0.31;
 Matches 83; Conservative 322; Mismatches 360; Indels 1; Gaps 1;

QY 839 TTATCCTCAACATTTAAACCTGGATCTATGAAATCAACAGCTGGTAAATTCGA 898
 DB 291 KKSYSRWCMTYGGGCGRATRTYWGYSRPAAMTKKMTYRGYGMKGMMAGMMNR 350

QY 899 CTAGACGACATCACTACTAGTAATGACAGAAAGACATTCATGATGGCT 958
 DB 351 SMCWRSACACYMMWRMBRMBRBRMAKSSRSRBRKRCMKRKYKMBRYSRBRSC 410

QY 959 ATATATAGATATAGTGAAGAGTCTCTGATTTAGCTACTTAAAGATGAGCA 1018
 DB 411 RAAMRRCRSSGRMKXKCGCMTCRMKSYGMKRWKSKMAKRYKMSMYMRKKKCSR 470

QY 1019 CACGAATGAGAAATATATCATGCAATTTCTGTAGCTACGACATGCAAGATTCTCA 1078
 DB 471 TTMGKTIRGMMOTMRCRYKRSKMKCRRRRMBRMYRMYRMYRMYRMYRMYR 530

QY 1079 AACTGACCGAGTTTCCCAACTACTGTGATGTTCACTGATCTTCACTGCTT 1138
 DB 531 KYSYSARARACMYRSGKYMMAGMMKRYKRYMMYMMYMMYMMYMMYMMYMMY 590

QY 1139 ATTATGTG-GAGTAGAATAGATTTCCACCAATTGAAATGACAAAGCAGAAATTTGCT 1197
 DB 591 MKSARAGAMCKRSMASMSKMSRSRCKRCASRSRASKRYAMMGCTSGSRMSRWS 650

QY 1198 TTATCTGTGGTAAATAGCTTTTCTCCAGTTGTATAAAGACCTCCACACGATATA 1257
 DB 651 YTCYWRKMSKSTCTMYYMSKYTAKGYSYMRRYRAMCMYMRMYRYRYRYMYTVA 710

QY 1258 GTCTATGCAACAAAGAAATGCAATCATCTCTTACTCTTATTTATTTCAATG 1317
 DB 711 WYTSSTRAMTGMKSGRYTSMYKCKMSYKRSBMYWYMMWMAKTMWRRYATRMW 770

QY 1318 ATAGCCGTTTCTTACTACACTCAATTAAGTAAGCAAGATGAGTGTGACTGT 1377
 DB 771 MMYRYSMKRYTCTMGGYMMYMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 830

QY 1378 TTATAAAGAAAGATAAAGATATCATCATTTAGGCAATTAAGGAGAGAGATT 1437
 DB 831 KMRMGWTAKTGRARARAYMYKMATWCACTKRMWTKGAGAKWATWMAKAYMYWNR 890

QY 1438 CAGCAACAGTGTCTTACAAGTGAAGAAAGAAAGTTAAAGTGAAGCCCTCTTGA 1497
 DB 891 AMYYYKTRRTYKTCMMKARWGSWAHYRMMWKGSAKMMWMMKGRMOWTKYMYVCTTW 950

QY 1498 GAAGATCAATGCAAGTTGACTTATGACGACCATCATCATCATTAATGCTTCG 1557
 DB 951 KMAAGATATYMCACGMAAMYSYSWTRTYMMKTRMMMASSRTAKMAAMMMKTRAMSXY 1010

CC from wipo.int/pub/published_pct_sequences
XX Sequence 2636 BP; 849 A; 414 C; 497 G; 876 T; 0 U; 0 Other;
SQ

Query Match 2.0%; Score 47; DB 4; Length 2636;
Best Local Similarity 52.3%; Pred. No. 1.1;
Matches 104; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 1821 ATTGCCCTGCACCTATTGTAAGAACTCTATAGAGGGAATACAAAGAAACAACTT 1880
DB 540 ATTGACTTCGAAATACCTTTAAATACAGCCCTGTCGTTCTTCGAAATTAACAT 599
QY 1881 CTGATTTTATTGCACTTCGATATCTTATGCAATTAGCTAATTCAGTAGAGCATTC 1940
DB 600 ATAAATTAATTTAAATGCTGAATTAATTCCTCTAAATAATCAGGTATATTATTTC 659
QY 1941 AGCAGAAATTTAAATAGAAATTAATAGTAAGAAATTTATTGATAGACTCTTGAAA 2000
DB 660 TTAGGGAAATTAAGTATTATTGCTGCTGTTTATTATGTTAAATAAGTATAGACAA 719
QY 2001 ATTACACAAGAGGAAATT 2019
DB 720 ATTAGTAGATGTGTAATTT 738

RESULT 14
ADD29587
ID ADD29587 standard; mRNA; 662 BP.
XX
AC ADD29587;
XX
DT 15-JAN-2004 (first entry)
XX
DE Mouse tumour suppressor mRNA SEQ ID NO:36.
XX
KM ss; mouse; tumour suppressor; cancer; cancer; cytostatic; gene therapy.
XX
OS Mus musculus.
XX
PN M02003058201-A2.
XX
PD 17-JUL-2003.
XX
PF 31-DEC-2002; 2002WO-US041825.
XX
PR 31-DEC-2001; 2001US-0345317P.
XX
PA (QUAR-) QUARK BIOTRCH INC.
XX
PA (CLEV-) CLEVELAND CLINIC FOUND.
XX
PI Feinstein E, Gudkov AV;
XX
DR WPI; 2003-598393/56.
XX
PT Diagnosing cancer comprises determining the polypeptide or polynucleotide
PT level e.g., hepatic lipase, in a sample from a subject, where a higher
PT level compared to that in a subject free of cancer is indicative of
PT cancer.
XX
PS Disclosure; SEQ ID NO 36; 272bp; English.
XX
XX The invention relates to a novel method for diagnosing a cancer in a
CC subject. the method comprises determining, in a sample from the cancer in a
CC the level of at least one polypeptide, where a higher level of the
CC polypeptide compared to the level of the polypeptide in a subject free of
CC cancer is indicative of cancer. The polypeptide is selected from any of
CC the polypeptides encoded by the polynucleotides listed in the
CC specification and polypeptides which are at least 70% homologous to the
CC polypeptides. The method of the invention has cytostatic activity, and
CC may have a use in gene therapy. The method is useful in identifying
CC markers specific for one or several types of cancer, depending on the
CC tissue origin, which may be used in numerous diagnostic and prognostic
CC applications as well as cancer type-specific targets for therapeutic

CC intervention. The compounds that modulate the activity of a tumour
CC suppressor gene are useful in the treatment of cancer or as anti-cancer
CC drugs. The present sequence represents a polynucleotide of the invention.
XX
SQ Sequence 662 BP; 217 A; 133 C; 148 G; 164 T; 0 U; 0 Other;
XX

Query Match 2.0%; Score 46.8; DB 10; Length 662;
Best Local Similarity 73.2%; Pred. No. 0.8;
Matches 60; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1615 ACAGCAGAAAGATGCGATTTAAGCGTACTTGAAAAAATGAGAAATGAGAACTTATGA 1674
DB 28 ACAGCTGAGATCAGGCACTTCAGCGCACGTGAAAGTAGACCCGAAACGAAACTATGAA 87
QY 1675 AAATTCATGAGCAATGCGTA 1696
DB 88 AAGTTCATGAGCAAAATGCGCA 109

RESULT 15
ADZ62521
ID ADZ62521 standard; DNA; 670 BP.
XX
AC ADZ62521;
XX
DT 30-JUN-2005 (first entry)
XX
DE Murine Fabp2 gene, SEQ ID 2245.
XX
KM Drug screening; Peroxisome Proliferator-Activated Receptor alpha;
KM PPAR alpha; gene; db.
XX
OS Mus musculus.
XX
PN US2005084872-A1.
XX
PD 21-APR-2005.
XX
PF 23-JAN-2004; 2004US-00764420.
XX
PR 24-JAN-2003; 2003US-0442797P.
XX
PR 30-MAY-2003; 2003US-0474413P.
XX
PA (LUMP/) LUM P Y.
PA (TANY/) TAN Y.
PA (DAIH/) DAI H.
PA (MUIS/) MUISE E S.
PA (BERG/) BERGER J P.
PA (THOM/) THOMPSON J R.
XX
PI Lum PY, Tan Y, Dai H, Muise ES, Berger JP, Thompson JR;
XX
DR WPI; 2005-313963/32.
XX
DR REFSEQ; NM_007980.
XX
PT Determining whether agent has defined biological activity, by making
PT comparison e.g. comparing efficacy value, toxicity value and classifier
PT value of agent to reference values, and using results to determine
PT whether agent has activity.
XX
PS Example 4; SEQ ID NO 2245; 51bp; English.
XX
XX The present invention relates to a method for screening biologically
CC active agents, such as candidate drugs, to determine whether the agent
CC possesses a defined biological activity. The methods involve making one
CC or more comparisons chosen from comparing efficacy value of agent to
CC reference efficacy value, comparing a toxicity value of the agent to
CC reference toxicity value, comparing a classifier value of the agent to
CC reference classifier value, and using the comparison result(s) to
CC determine whether the agent possesses the defined biological activity.
CC Also claimed is a population of oligonucleotide probes (I) specific for
CC measuring the expression levels of members of a classifier population of
CC genes or a toxicity-related population of genes. (I) are useful for

CC measuring the expression levels of genes that are useful for identifying
 CC agonists or partial agonists of Peroxisome Proliferator-Activated
 CC Receptor (PPAR) gamma. ADZ63072-ADZ63059 are oligonucleotide probes which
 CC are useful in the method of the invention to measure the expression
 CC pattern of mouse genes ADZ6273-ADZ63071 that yield an expression pattern
 CC that correlates with the stimulation of PPARalpha receptors by an agent.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html?DocID=20050084872.

XX
 SQ Sequence 670 BP; 228 A; 131 C; 147 G; 164 T; 0 U; 0 Other;

Query Match 2.0%; Score 46.8; DB 14; Length 670;
 Best Local Similarity 73.2%; Pred. No. 0.81;
 Matches 60; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1615 ACAGACAGAAAGATGGCATTTAACGTACTTGGAAAAATGAGAAATGAACTATGAA 1674
 Db 24 ACAGCTGAGATCATGCGATTGACGGCAGTGGAAAGTAAGCCGACGAGAACTATGAA 83

QY 1675 AAATTGATGGAAGCAATGGGTA 1696
 Db 84 AAGTTGATGGAAGAAATGGGCA 105

Search completed: May 27, 2006, 19:15:54
 Job time : 1601.66 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 28, 2006, 05:53:57 ; Search time 12207.9 Seconds
(without alignments)
10906.400 Million cell updates/sec

Title: US-10-099-663-1

Perfect score: 2381

Sequence: 1 agctccctgcgcagcaaaag.....atctgaagctcactatcag 2381

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_est7:*
7: gb_est8:*
8: gb_est9:*
9: gb_est10:*
10: gb_est11:*
11: gb_est12:*
12: gb_est13:*
13: gb_est14:*
14: gb_est15:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	752.6	31.6	1473	12	CC250935 CH261-17C
2	99	4.2	598	5	CF250818 esa017_e0
3	95	4.0	595	5	CF250947 esa019_a0
4	85	3.6	705	3	BU255106 60347428
5	84	3.5	759	3	BU265304 603508640
6	81	3.4	829	3	BU296697 603741965
7	79.4	3.3	885	3	BU212380 603003274
8	79	3.3	814	3	BU234099 603792674
9	78.8	3.3	1080	3	BU212336 603149475
10	75.8	3.2	426	5	CD739695 4028031_1
11	74	3.1	576	5	CD739141 4028481_1
12	70.4	3.0	480	5	CD733346 4045132_1
13	62.4	2.6	1101	14	CNS00396
14	61.4	2.6	885	3	BU123717 603147391
15	60.4	2.5	443	4	BK276255 BX276255
16	57.2	2.4	683	8	CW793516 4128553_B
17	55.8	2.3	657	5	CK945920 4070423_B
18	55	2.3	730	10	DT824420 LB00233.C
19	55	2.3	731	10	DT824909 LB00235.C

20	54.2	2.3	572	5	CK834687
21	54.2	2.3	673	5	CK947902
22	54.2	2.3	686	5	CK957350
23	54.2	2.3	708	5	CK960577
24	54	2.3	481	2	BM430667
25	54	2.3	489	4	CB224035
26	54	2.3	497	2	BM430956
27	54	2.3	548	2	BM432416
28	54	2.3	568	2	BM431104
29	54	2.3	671	8	CK980160
30	54	2.3	679	5	CK947198
31	54	2.3	682	5	CK948900
32	54	2.3	684	5	CK947048
33	54	2.3	702	5	CK944155
34	54	2.3	720	5	CK946512
35	53.4	2.2	669	10	DT820486
36	53.4	2.2	726	10	DT820499
37	53.4	2.2	726	10	DT821139
38	53.4	2.2	726	10	DT822486
39	53.4	2.2	726	10	DT825113
40	53.4	2.2	727	10	DT819495
41	53.4	2.2	727	10	DT823340
42	53.4	2.2	823	10	DT822459
43	53.4	2.2	843	10	DT828569
44	53	2.2	372	2	BM430450
45	53	2.2	414	2	BM433066

ALIGNMENTS

RESULT 1
LOCUS CC250935/1473 bp DNA linear GSS 13-MAY-2003
DEFINITION CH261-17C7_Sp6.1 CH261 Gallus gallus genomic clone CH261-17C7,
genomic survey sequence.

ACCESSION CC250935
VERSION CC250935.1 GI:30587685
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 1473)
AUTHORS Kremetzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
Archosouria,Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

TITLE Warren,W., Graves,T., Mardis,E. and Wilson,R.
JOURNAL Gallus gallus BAC End Reads
COMMENT Unpublished (2003)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTACGTGACACTATG
Class: BAC ends
High quality sequence start: 239
High quality sequence stop: 912.
Location/Qualifiers

FEATURES

source

1..1473

/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-17C7"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_id="CH261"
/note="Vector: pPARAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - For library and clone
ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match	31.6%	Score 752.6	DB 12	Length 1473	
Best Local Similarity	98.3%	Pred. No. 6,2e-168			
Matches 793	Conservative 0	Mismatch 9	Indels 5	Gaps 3	
Qy	1580	TTTGCAGAACTCTGCTACTTACCGAGAAATC-TGCCTACAGACAGAAAGATGGCATTTTAA- TTGCAGAAAGTTGCTGCTTCTTACAGAAATCTTCCCTACAGACAGAAAGATGGCATTTTAA	163	163	
Db	1042	TTGCAGAAAGTTGCTGCTTCTTACAGAAATCTTCCCTACAGACAGAAAGATGGCATTTTAA	983	983	
Qy	1638	CGGTACTTGGGAAATGAGAAAAATGAGAACTTATGAAAAATTCATGAAAGCATGGCTTAA CGGTACTTGGGAAATGAGAAAAATGAGAACTTATGAAAAATTCATGAAAGCATGGCTTAA	1697	1697	
Db	982	CGGTACTTGGGAAATGAGAAAAATGAGAACTTATGAAAAATTCATGAAAGCATGGCTTAA	923	923	
Qy	1698	GCCTTACTTTTGTGAATGCTTCTAATAACAGAGATTCACCTACGGCGGAAATACAACTTAA GCCTTACTTTTGTGAATGCTTCTAATAACAGAGATTCACCTACGGCGGAAATACAACTTAA	1757	1757	
Db	922	GCCTTACTTTTGTGAATGCTTCTAATAACAGAGATTCACCTACGGCGGAAATACAACTTAA	863	863	
Qy	1758	AGCTGTTCAATGAATCAATCTGGCTAACCTGTCTCTTGTGTCTATATTTGGCCCTTG AGCTGTTCAATGAATCAATCTGGCTAACCTGTCTCTTGTGTCTATATTTGGCCCTTG	1817	1817	
Db	862	AGCTGTTCAATGAATCAATCTGGCTAACCTGTCTCTTGTGTCTATATTTGGCCCTTG	803	803	
Qy	1818	CACATTTGCCCTGACCTTATTTGAAAAAGCTGTATATGAGGGGAAATACAAAGAAAGAAAC CACATTTGCCCTGACCTTATTTGAAAAAGCTGTATATGAGGGGAAATACAAAGAAAGAAAC	1877	1877	
Db	802	CACATTTGCCCTGACCTTATTTGAAAAAGCTGTATATGAGGGGAAATACAAAGAAAGAAAC	743	743	
Qy	1878	ATTCTGATTTTATTTGCAATTTGGCATTTGCGATATCTTATGCAATTTAGCTAATTCAGTAGAGCAT ATTCTGATTTTATTTGCAATTTGGCATTTGCGATATCTTATGCAATTTAGCTAATTCAGTAGAGCAT	1937	1937	
Db	742	ATTCTGATTTTATTTGCAATTTGGCATTTGCGATATCTTATGCAATTTAGCTAATTCAGTAGAGCAT	683	683	
Qy	1938	TCCACAGAAATTTTAAATGAAATTTATATGTAAGAAATTTTATTTTGAATAAACATGTTTGA TCCACAGAAATTTTAAATGAAATTTATATGTAAGAAATTTATTTTGAATAAACATGTTTGA	1997	1997	
Db	682	TCCACAGAAATTTTAAATGAAATTTATATGTAAGAAATTTATTTTGAATAAACATGTTTGA	623	623	
Qy	1998	AAAAATTACACAAGAGGAAATTTGCTGCTCCAGCTTTTGACAGAACACATGATTTTGAGT AAAAATTACACAAGAGGAAATTTGCTGCTCCAGCTTTTGACAGAACACATGATTTTGAGT	2057	2057	
Db	622	AAAAATTACACAAGAGGAAATTTGCTGCTCCAGCTTTTGACAGAACACATGATTTTGAGT	563	563	
Qy	2058	CATTTTAAACATGCTAGTGTCTTAAAGCTTGTACAACTGCTGTATATGATGTAA CATTTTAAACATGCTAGTGTCTTAAAGCTTGTACAACTGCTGTATATGATGTAA	2117	2117	
Db	562	CATTTTAAACATGCTAGTGTCTTAAAGCTTGTACAACTGCTGTATATGATGTAA	503	503	
Qy	2118	ACATTAATCTATCTAGTGTGATAGTAGTTGTATTAACAGGCTGAAACATGCTCAGTGTAA ACATTAATCTATCTAGTGTGATAGTAGTTGTATTAACAGGCTGAAACATGCTCAGTGTAA	2177	2177	
Db	502	ACATTAATCTATCTAGTGTGATAGTAGTTGTATTAACAGGCTGAAACATGCTCAGTGTAA	443	443	
Qy	2178	GGTGAGAAAGATTAAGACTCTGAGTCAGAAATTTGGGCTAAGCTCCCTGAACCTACAGAAA GGTGAGAAAGATTAAGACTCTGAGTCAGAAATTTGGGCTAAGCTCCCTGAACCTACAGAAA	2237	2237	
Db	442	GGTGAGAAAGATTAAGACTCTGAGTCAGAAATTTGGGCTAAGCTCCCTGAACCTACAGAAA	383	383	
Qy	2238	AAGTCACATATAAAAATGCAAAATGATGATCTATTTTGTCTCTGCTGTGATGTAA AAGTCACATATAAAAATGCAAAATGATGATCTATTTTGTCTCTGCTGTGATGTAA	2297	2297	
Db	382	AAGTCACATATAAAAATGCAAAATGATGATCTATTTTGTCTCTGCTGTGATGTAA	323	323	
Qy	2298	TGATTTATTTTAA--TTTGTGTGTGTGTAGCGCTGATGTGATGAAAGAAATTAGAGCC TGATTTATTTTAA--TTTGTGTGTGTGTAGCGCTGATGTGATGAAAGAAATTAGAGCC	2354	2354	
Db	322	TGATTTATTTTAA--TTTGTGTGTGTGTAGCGCTGATGTGATGAAAGAAATTAGAGCC	263	263	
Qy	2355	CACGATATCTGAAGCTACATATTCAG 2381 CACGATATCTGAAGCTACATATTCAG 2381	2381	2381	
Db	262	CACGATATCTGAAGCTACATATTCAG 2381	2381	2381	
RESULT 2	CF250818	598 bp	mRNA	linear	EST 07-AUG-2003
LOCUS	CF250818				
DEFINITION	ese017_m95 Elmeiriria temella-infected caecal tonsil Gallus gallus				
ACCESSION	CF250818				
VERSION	CF250818.1				
KEYWORDS	EST.				
SOURCE	Gallus gallus (chicken)				
ORGANISM	Gallus gallus				
	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archaeopteryx; Aves; Neognathae; Galliformes; Phasianidae;				

```

REFERENCE          Phaenianinae; Gallus.
AUTHORS            1 (bases 1 to 598)
                   Wiltzell,H., Bed'Hom,B., Morin,V., Young,J.R., Whitaker,C.J.,
                   Chausse,A.M., and Zoorob,R.
TITLE              A collection of chicken ESTs from activated immune cells
JOURNAL            Unpublished (2003)
COMMENT            Contact: Zoorob R
                   UPR 1983
CNRS
7 rue Guy Moquet, BP 8, 94801 Villejuif cedex, France
Tel: 33 1 49 58 35 00
Fax: 33 1 49 58 33 81
Email: zoorob@vjf.cnrs.fr.
Location/Qualifiers
1. .598
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone_lib="Bimetricia tenella-infected caecal tonsil"
/notes="Organ: Caecal tonsil; Vector: pTribEX2"

ORIGIN
Query Match          4.2%; Score 99; DB 5; Length 598;
Best Local Similarity 100.0%; Pred. No. 2.7e-12;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1596  ACTTACCGAAGTCTGCCTACAGACAGAAAGATGCAATTAAACGCTACTTGGAAAAATGGA 16550303
Db      3      ACTTACCGAAGTCTGCCTACAGACAGAAAGATGCAATTAAACGCTACTTGGAAAAATGGA 62

Oy      1656  GAAAAATGAGAACTATGAAAAAATTCATGAGACATGGG 1694
Db      63      GAAAAATGAGAACTATGAAAAAATTCATGAGACATGGG 101

RESULT 3
LOCUS      CF250947                    595 bp      mRNA      linear      EST 07-AUG-2003
DEFINITION eae019 a07 Bimetricia tenella-infected caecal tonsil Gallus gallus
            cDNA, mRNA sequence.
ACCESSION  CF250947
VERSION    CF250947.1  GI:33484202
KEYWORDS   EST.
SOURCE     Gallus gallus (chicken)
            Gallus gallus
            Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Eucleostomi;
            Archosauilia; Aves; Neognathae; Galliformes; Phaenianidae;
            Phaenianinae; Gallus.
            1 (bases 1 to 595)
REFERENCE  Wiltzell,H., Bed'Hom,B., Morin,V., Young,J.R., Whitaker,C.J.,
            Chausse,A.M., and Zoorob,R.
            A collection of chicken ESTs from activated immune cells
            Unpublished (2003)
            Contact: Zoorob R
            UPR 1983
CNRS
7 rue Guy Moquet, BP 8, 94801 Villejuif cedex, France
Tel: 33 1 49 58 35 00
Fax: 33 1 49 58 33 81
Email: zoorob@vjf.cnrs.fr.
Location/Qualifiers
1. .595
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone_lib="Bimetricia tenella-infected caecal tonsil"
/notes="Organ: Caecal tonsil; Vector: pTribEX2"

ORIGIN
Query Match          4.0%; Score 95; DB 5; Length 595;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1600 ACCAGAGCTCTGCTACAGACAGAAAGATGCGATTAAACGTAATTGAGAAATAGAGAA 1659
 DB 3 ACCAGAGCTCTGCTACAGACAGAAAGATGCGATTAAACGTAATTGAGAAATAGAGAA 62
 QY 1660 AATGGAATCTAGAAATTTCTGAGAGCAATGGG 1694
 DB 63 AATGGAATCTAGAAATTTCTGAGAGCAATGGG 97

RESULT 4
 LOCUS BU355106 705 bp mRNA linear EST 28-NOV-2002
 DEFINITION 603474288P1 CSEQCHN70 Gallus gallus cDNA clone CHEST355022 5', mRNA sequence.

ACCESSION BU355106
 VERSION BU355106.1 GI:25863107
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Phasianidae; Aves; Neognathae; Galliformes; Phasianidae; Gallus.

REFERENCE 1 (bases 1 to 705)
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)

COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES
 source location/Qualifiers

1..705
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, Hisex"
 /db_xref="taxon:9031"
 /clone="CHEST355022"
 /dev_stage="36"
 /lab_host="DH10B"
 /clone_lib="CSEQCHN70"
 /note="Organ: hearts; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was bluntended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Query Match 3.6%; Score 85; DB 3; Length 705;
 Best Local Similarity 100.0%; Pred. No. 6.2e-09;
 Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1610 TGCCCTACAGACAGAAAGATGCGATTAAACGTAATTGAGAAATAGAGAAATAGAGAACT 1669
 DB 1 TGCCCTACAGACAGAAAGATGCGATTAAACGTAATTGAGAAATAGAGAAATAGAGAACT 60
 QY 1670 ATGAAATTTCTAGAGAGCAATGGG 1694
 DB 61 ATGAAATTTCTAGAGAGCAATGGG 85

RESULT 5
 LOCUS BU265304 759 bp mRNA linear EST 26-NOV-2002
 DEFINITION 603508640P1 CSEQCHN52 Gallus gallus cDNA clone CHEST437010 5', mRNA sequence.

ACCESSION BU265304
 VERSION BU265304.1 GI:25536254
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Phasianidae; Aves; Neognathae; Galliformes; Phasianidae; Gallus.

REFERENCE 1 (bases 1 to 759)
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)

COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES
 source location/Qualifiers

1..759
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, Hisex"
 /db_xref="taxon:9031"
 /clone="CHEST437010"
 /dev_stage="22"
 /lab_host="DH10B"
 /clone_lib="CSEQCHN52"
 /note="Organ: limbs; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was bluntended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Query Match 3.5%; Score 84; DB 3; Length 759;
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1611 GCCTACAGACAGAAAGATGCGATTAAACGTAATTGAGAAATAGAGAAATAGAGAACTA 1670
 DB 1 GCCTACAGACAGAAAGATGCGATTAAACGTAATTGAGAAATAGAGAAATAGAGAACTA 60
 QY 1671 TGAATTTCTAGAGAGCAATGGG 1694
 DB 61 TGAATTTCTAGAGAGCAATGGG 84
 RESULT 6
 LOCUS BU296697 829 bp mRNA linear EST 27-NOV-2002
 DEFINITION 603741965P1 CSEQCHN56 Gallus gallus cDNA clone CHEST640b17 5', mRNA sequence.

ACCESSION BU296697
 VERSION BU296697.1 GI:25746333
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus; 1 (bases 1 to 829)
 REFERENCE Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J., A Comprehensive Collection of Chicken cDNAs Curr. Biol. 12 (22), 1965-1969 (2002)
 AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
 TITLE A Comprehensive Collection of Chicken cDNAs
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
 PUBMED 12445392
 COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers
 1..829
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Compton Line 151"
 /db_xref="taxon:9031"
 /clone="CHEST640b17"
 /sex="Female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_1lb="CSEQCHN56"
 /note="Organ: small intestine; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunt-ended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1996): 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN
 Query Match 3.4%; Score 81; DB 3; Length 829;
 Best Local Similarity 100.0%; Pred. No. 5,7e-08;
 Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1614 TACAGACGAAAGATGGCATTTAAACGTAAGTGAAGAAATGAGAACTATGA 1673
 |||||||
 DB 11 TACAGACGAAAGATGGCATTTAAACGTAAGTGAAGAAATGAGAACTATGA 70
 QY 1674 AAAATTGATGAGCAATGGG 1694
 |||||||
 DB 71 AAAATTGATGAGCAATGGG 91
 RESULT 7
 BU123280 885 bp mRNA linear EST 25-NOV-2002
 LOCUS 603003274F1 CSEQCHL18 Gallus gallus cDNA clone CHEST14C20 5', mRNA
 DEFINITION sequence.
 ACCESSION BU123280
 VERSION BU123280.1 GI:25333903
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus; 1 (bases 1 to 885)
 AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J., A Comprehensive Collection of Chicken cDNAs Curr. Biol. 12 (22), 1965-1969 (2002)
 TITLE A Comprehensive Collection of Chicken cDNAs
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
 PUBMED 12445392
 COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers
 1..885
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Compton Line 151"
 /db_xref="taxon:9031"
 /clone="CHEST14C20"
 /sex="Female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_1lb="CSEQCHL18"
 /note="Organ: small intestine; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; Modification of pBluescript II KS(+) (Stratagene) vector to accommodate cDNA produced with the T-trimmed protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3' end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI. Ligate in double stranded adaptor containing BspGI and BamHI sites [5'ggccgcgtcagcccgccgagtcggaataaag] [5'aattcttttttcggatccggcgccgcgc]".

ORIGIN
 Query Match 3.3%; Score 79.4; DB 3; Length 885;
 Best Local Similarity 98.8%; Pred. No. 1.4e-07;
 Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1614 TACAGACGAAAGATGGCATTTAAACGTAAGTGAAGAAATGAGAACTATGA 1673
 |||||||
 DB 1 TACAGACGAAAGATGGCATTTAAACGTAAGTGAAGAAATGAGAACTATGA 60
 QY 1674 AAAATTGATGAGCAATGGG 1694
 |||||||
 DB 61 AAAATTGATGAGCAATGGG 81
 RESULT 8
 BU234099 814 bp mRNA linear EST 26-NOV-2002
 LOCUS 603792609P1 CSEQCHN24 Gallus gallus cDNA clone CHEST758b4 5', mRNA
 DEFINITION sequence.
 ACCESSION BU234099
 VERSION BU234099.1 GI:25478348
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus; 1 (bases 1 to 814)
 REFERENCE Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J., A Comprehensive Collection of Chicken cDNAs Curr. Biol. 12 (22), 1965-1969 (2002)
 TITLE A Comprehensive Collection of Chicken cDNAs
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
 PUBMED 12445392
 COMMENT Contact: Simon Hubbard

/strain="white leghorn SC"
 /db_xref="taxon:9031"
 /clone="IGAL_20802"
 /sex="mixed"
 /tissue_type="Gut"
 /cell_type="Lymphocyte"
 /dev_stage="Adult"
 /lab_host="EMDH10B"
 /clone_lib="IGAL - Chicken Intestinal Lymphocyte"
 /note="Organ: Intestine; Vector: PCMV-SPORT6; Site 1:
 Salt; Site 2: NoCI; Normalized library from chicken gut
 infected with coccidia duodenum and middle gut."

ORIGIN

Query Match 3.2%; Score 75.8; DB 5; Length 426;
 Best Local Similarity 97.5%; Pred. No. 8.5e-07;
 Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1616 CAGCAGAAAGATGGCGTTTAAACGCTACTTGAAGAAATAGAGAAATGAGAACTATGAAA 1675
 |||||
 Db 2 CAGCAGAAAGATGGCGTTTAAACGCTACTTGAAGAAATAGAGAAATGAGAACTATGAAA 61
 |||||
 QY 1676 AATTCATGGAAGCAATGGG 1694
 |||||
 Db 62 AATTCATGGAAGCAATGGG 80
 |||||

RESULT 11
 CD739141 576 bp mRNA linear EST 26-JUN-2003
 LOCUS CD739141
 DEFINITION 4026481 IGAL - Chicken Intestinal Lymphocyte Gallus gallus CDNA
 clone IGAL_93A12 5', mRNA sequence.
 ACCESSION CD739141 GI:32289990
 VERSION CD739141.1 GI:32289990
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianine; Gallus.
 1 (bases 1 to 576)
 Min.W., Lillehoj,H.S., Ashwell,C.M., Matukumalli,L.K., van
 Tassel,C. and Han,J.Y.
 Chicken Intestinal Lymphocyte EST database as a resource for the
 analysis of mucosal immune function
 Unpublished (2003)
 Contact: Hyun S. Lillehoj
 Animal Parasite Diseases Laboratory
 Animal and Natural Resources Institute, USDA
 Bldg.1043, BARC-East, Beltsville, MD 20705, USA
 Tel: 3015048771
 Fax: 3015045103
 Email: lillehoj@nri.barc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred
 0.000925 using options -trim_alt '-trim_faasta. Vector identified
 by cross_match using options -mismatch 12 -mismatch 12
 Plate: 93 row: A column: 12
 Seq primer: ATTTAGGTGACACTATAG
 High quality sequence stop: 576.
 Location/Qualifiers
 1..576

FEATURES

source

/organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="white leghorn SC"
 /db_xref="taxon:9031"
 /clone="IGAL_93A12"
 /sex="mixed"
 /tissue_type="Gut"
 /cell_type="Lymphocyte"
 /dev_stage="Adult"
 /lab_host="EMDH10B"
 /clone_lib="IGAL - Chicken Intestinal Lymphocyte"
 /note="Organ: Intestine; Vector: PCMV-SPORT6; Site_1:
 infected with coccidia duodenum and middle gut."

Salt; Site 2: NoCI; Normalized library from chicken gut
 infected with coccidia duodenum and middle gut."

ORIGIN

Query Match 3.1%; Score 74; DB 5; Length 576;
 Best Local Similarity 100.0%; Pred. No. 2.5e-06;
 Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1621 AGAAAGATGGCATTTAAACGCTACTTGAAGAAATAGAGAAATGAGAACTATGAAAATTC 1680
 |||||
 Db 1 AGAAAGATGGCATTTAAACGCTACTTGAAGAAATAGAGAAATGAGAACTATGAAAATTC 60
 |||||
 QY 1681 ATGGAAGCAATGGG 1694
 |||||
 Db 61 ATGGAAGCAATGGG 74
 |||||

RESULT 12
 CD733346 480 bp mRNA linear EST 26-JUN-2003
 LOCUS CD733346
 DEFINITION 4045132 IGAL - Chicken Intestinal Lymphocyte Gallus gallus CDNA
 clone IGAL_41E15 5', mRNA sequence.
 ACCESSION CD733346
 VERSION CD733346.1 GI:32284195
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianine; Gallus.
 1 (bases 1 to 480)
 Min.W., Lillehoj,H.S., Ashwell,C.M., Matukumalli,L.K., van
 Tassel,C. and Han,J.Y.
 Chicken Intestinal Lymphocyte EST database as a resource for the
 analysis of mucosal immune function
 Unpublished (2003)
 Contact: Hyun S. Lillehoj
 Animal Parasite Diseases Laboratory
 Animal and Natural Resources Institute, USDA
 Bldg.1043, BARC-East, Beltsville, MD 20705, USA
 Tel: 3015048771
 Fax: 3015045103
 Email: lillehoj@nri.barc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred
 0.000925 using options -trim_alt '-trim_faasta. Vector identified
 by cross match using options -mismatch 12 -mismatch 12
 Plate: 41 row: B column: 15
 Seq primer: ATTTAGGTGACACTATAG
 High quality sequence stop: 480.
 Location/Qualifiers
 1..480

FEATURES

source

/organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="white leghorn SC"
 /db_xref="taxon:9031"
 /clone="IGAL_41E15"
 /sex="mixed"
 /tissue_type="Gut"
 /cell_type="Lymphocyte"
 /dev_stage="Adult"
 /lab_host="EMDH10B"
 /clone_lib="IGAL - Chicken Intestinal Lymphocyte"
 /note="Organ: Intestine; Vector: PCMV-SPORT6; Site 1:
 Salt; Site 2: NoCI; Normalized library from chicken gut
 infected with coccidia duodenum and middle gut."

ORIGIN

Query Match 3.0%; Score 70.4; DB 5; Length 480;
 Best Local Similarity 98.6%; Pred. No. 1.7e-05;
 Matches 71; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1623 AAGATGGCATTTAAACGCTACTTGAAGAAATGAGAAATGAGAACTATGAAAATTCAT 1682
 |||||

D	b		745	DDRRBATTAGKRRRTTKRKWRKRDRDTWMDADBDATDARRRGGGCGAAGCKTGRK	686
Oy			2142	AGTTTGATTAACAGGCCTGAACA CTGCTCAGTCGTAAAGSTGAGAAGTAGACTGAC	2201
D	b		685	RRRDRATRWDTRTDMAWMADAAMWT TTDITD DMPDKDRRRKGARRRRTTBAAA DMWTWK	626
Oy			2202	TCA GATA TTCTGGGCTA AGCTCCCT CCA ACT CA GA AAAAAAGCAC ATATA AAATCAA CAT	2261
D	b		625	- AMDMA CKMT RADR WDA ADTW TD AR K DR MA RA RM RR AR DBRA RD RR WT X	567
Oy			2262	GAT GT TC ATA ATT TG TTT CTC T GC TC G TA GT GTA AT GA TTA TTA TTA TTT TTT TTA	2321
D	b		566	GKT TTA TWTT AAA BAAA MAMA NAWAT TTA TT TTT TTT TTT TTT TTT TTT TTA MMWA	507
Oy			2322	GG CGTGA TGTGA TGA AAAAA GGTT	2347
D	b		506	WWTATWAA WTTAA AAAAAAAAA MATT	481
RESULT 14	BUI23717				
LOCUS	BUI23717				
DEFINITION	BUI23717	885 bp	mRNA	linear	EST-25-NOV-2002
ACCESSION	60314739.F1	CSEQCHL18	Gallus gallus	cDNA clone CHEST14fE24 5'	mRNA sequence.
VERSION	BUI23717				
KEYWORDS	BUI23717.1	GI:25334366			
SOURCE	EST.				
ORGANISM	Gallus gallus (chicken)				
AUTHORS	Eukaryotic; Metazoa; Chordata; Craniata; Vertebrata; Eucleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae;				
JOURNAL PUBMED COMMENT	Phasianinae; Gallus. 1 (bases 1 to 885) Boardman,P.E., Sans-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,B., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J. A Comprehensive Collection of Chicken CDNAS Curr. Biol. 12 (22), 1965-1969 (2002) 12445392				
FEATURES SOURCE	Contact: Simon Hubbard Department of Biomolecular Sciences University of Manchester Institute of Science and Technology (UMIST) PO Box 88, Manchester, M60 1QD, UK Tel.: 01612008930 Fax: 01612360409 Email: Simon.Hubbard@umist.ac.uk. Location/Qualifiers 1..885 /organism="Gallus gallus" /mol_type="mRNA" /server="Compton Line 15I" /db_xref="taxon:9031" /clohes="CHEST14Ef24" /dev_stage="adult" /sex="Female" /lab_host="DH10B" /clone_1ib="CSEOCHL18" /note="Organ: small intestine; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; Modification of cDNA produced with the T-timed protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3' end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387,624). Out pbluescript II KS(+) with NotI and EcoRI. Ligate in double stranded adaptor containing BamHI sites [5'-ggcgccgtgcagcccgcgacatcgaagaagg] [5'-aatcttttcttcgatccgccggcgtagcacgc]"				
ORIGIN	Query Match	2.6%	Score 61.4;	DB 3;	Length 885;
	Best Local Similarity	96.1%;	Pred. No. 0.0027;		


```
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11761
; LENGTH: 8905
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11761
```

Query Match 2.1%; Score 49.8; DB 3; Length 8905;

Best Local Similarity 59.2%; Pred. No. 0.014; Indels 9; Gaps 2;

Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

```
QY 1516 TGAGCTTTAGCCAGCATCATCTGTAATTTGCTTCTGATAGCCTGTTCAATAAT 1575
DB 1944 TGAATTTTAGCTTCAACATCACTAGTATGAGTTGGTTCAAGATTAAGAAATTAAT 2003
QY 1576 TCTC-----TTGCAAAAGCTCTGCTACTTACCAAGAGTCTGCTTACAGAGAAAGA 1627
DB 2004 TCTCCGCCAGAGCAGACCTGATCTCTAGCTGCTAGAGGCTGACT-CAACTGAATCA 2062
QY 1628 TGGCATTAGCGTACTTGGAAATAGAGAAATAGAGACATCTGTAATTAATTCATGAG 1687
DB 2063 TGGCCTTTGACGACCTTGGAAAGTAGACCGGAGTGAATTAATGACAGTTTCATGAAA 2122
QY 1688 CAATGGTAAAGCCTTACTTTTGTGATGCT 1718
DB 2123 AAATGGTAAGACTTATTTCTTTGTGCT 2153
```

RESULT 4

```
US-09-949-016-16261
; Sequence 16261, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001107
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16261
; LENGTH: 8907
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16261
```

Query Match 2.1%; Score 49.8; DB 3; Length 8907;

Best Local Similarity 59.2%; Pred. No. 0.014; Indels 9; Gaps 2;

Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

```
QY 1516 TGAGCTTTAGCCAGCATCATCTGTAATTTGCTTCTGATAGCCTGTTCAATAAT 1575
DB 1944 TGAATTTTAGCTTCAACATCACTAGTATGAGTTGGTTCAAGATTAAGAAATTAAT 2003
QY 1576 TCTC-----TTGCAAAAGCTCTGCTACTTACCAAGAGTCTGCTTACAGAGAAAGA 1627
DB 2004 TCTCCGCCAGAGCAGACCTGATCTCTAGCTGCTAGAGGCTGACT-CAACTGAATCA 2062
```

```
QY 1628 TGGCATTAGCGTACTTGGAAATAGAGAAATAGAGACATCTGTAATTAATTCATGAG 1687
DB 2063 TGGCCTTTGACGACCTTGGAAAGTAGACCGGAGTGAATTAATGACAGTTTCATGAAA 2122
QY 1688 CAATGGTAAAGCCTTACTTTTGTGATGCT 1718
DB 2123 AAATGGTAAGACTTATTTCTTTGTGCT 2153
```

RESULT 5

```
US-09-621-976-2813/C
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813
```

Query Match 2.0%; Score 48.4; DB 3; Length 832;

Best Local Similarity 13.7%; Pred. No. 0.012; Indels 2; Gaps 2;

Matches 53; Conservative 180; Mismatches 151; Indels 2; Gaps 2;

```
QY 342 ACAATAGCAATCTAAGTTGTTCTTTCTCTCTCTTCTTCACTTCTGTAAGCTTAAGG 401
DB 384 ACCATTAATTAATTAATTTGTTCTTTCTCTCTCTTCTTCACTTCTGTAAGCTTAAGG 326
QY 402 TGAAGAGTAGCTATTGAGTACTTCCCTCTGCACTCTTACCGAGATTGATGATTT 461
DB 325 YACASRYRYKWTGMMWMMKMMSTRYCYMCKCMKRGRCAYTAAAGHMSYAMCK 266
QY 462 CAATTAAGCACTGAGTGAATGAAGCAAGCACTATTTTGTAATCAGCAAGATTCTA 521
DB 265 WKSMSASMSKCTRYKKGSTWTWKTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 207
QY 522 AATTATAGTATTAATCTGATTAACCTTTTCTGCTGAGCTGTAAGAAAGAAAGATT 581
DB 206 YTSRYSYGMVYASWYTWCMWMMGRWSTYVYAMGKMYRYATTMRBAMMMWAAWTMM 147
QY 582 TGATTAACCAAGACTAGTAATAATTCATTAGTTAGGCAACCTGTATCTGTGATTA 641
DB 146 YMMWAMCMSSRGAAYRRTMMWGYRWRRKSYRTRCAVAYAKTKRSYVWCWIKW 87
QY 642 GCAACATTCATTTCAGCATTCAGATTTTACATTTTGAAGCTAATAGACAGAGATTGG 701
DB 86 RCMWMMWMMWMAVYKTMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMM 27
QY 702 TGCCGTCATAGGAACAGACTAATA 727
DB 26 WKSMSWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMM 1
```

RESULT 6

```
US-09-385-982-474
; Sequence 474, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS: II
```

```
/ FILE REFERENCE: CCDNA-260XX
/ CURRENT APPLICATION NUMBER: US/09/385,982
/ CURRENT FILING DATE: 1999-08-30
/ EARLIER APPLICATION NUMBER: 09/328,111
/ EARLIER FILING DATE: 1999-06-08
/ EARLIER APPLICATION NUMBER: 60/117,393
/ EARLIER FILING DATE: 1999-01-27
/ EARLIER APPLICATION NUMBER: 60/098,639
/ EARLIER FILING DATE: 1998-08-31
/ NUMBER OF SEQ ID NOS: 544
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 474
/ LENGTH: 515
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)-(515)
/ OTHER INFORMATION: n = A,T,C or G
US-09-385-982-474
```

```
Query Match 1.9%; Score 44.2; DB 3; Length 515;
Best Local Similarity 75.3%; Pred. No. 0.12;
Matches 55; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 1623 AAAGATGGCATTTAAGCGTACTTGAAATAAGAGAAAATGAGAACTATGAAAAATTCAT 1682
DB 23 AATCATGGCGTTTGACAGCAGCTTGAAGGTAGACCGAGTGAATACTATGACAGTTCTAT 82
QY 1683 GGAAGCATGGGT 1695
DB 83 GGAATAAATGGGT 95
```

```
RESULT 7
US-09-949-016-4519
/ Sequence 4519, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4519
/ LENGTH: 2252
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-4519
```

```
Query Match 1.9%; Score 44.2; DB 3; Length 2252;
Best Local Similarity 75.3%; Pred. No. 0.23;
Matches 55; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
```

QY 1623 AAAGATGGCATTTAAGCGTACTTGAAATAAGAGAAAATGAGAACTATGAAAAATTCAT 1682
DB 58 AATCATGGCGTTTGACAGCAGCTTGAAGGTAGACCGAGTGAATACTATGACAGTTCTAT 117
QY 1683 GGAAGCATGGGT 1695
DB 118 GGAATAAATGGGT 130

RESULT 8

```
US-09-949-016-19
/ Sequence 19, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 19
/ LENGTH: 2273
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-19
```

```
Query Match 1.9%; Score 44.2; DB 3; Length 2273;
Best Local Similarity 75.3%; Pred. No. 0.23;
Matches 55; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 1623 AAAGATGGCATTTAAGCGTACTTGAAATAAGAGAAAATGAGAACTATGAAAAATTCAT 1682
DB 58 AATCATGGCGTTTGACAGCAGCTTGAAGGTAGACCGAGTGAATACTATGACAGTTCTAT 117
QY 1683 GGAAGCATGGGT 1695
DB 118 GGAATAAATGGGT 130
```

```
RESULT 9
US-08-232-463-14/C
/ Sequence 14, Application US/08232463
/ Patent No. 5670367
/ GENERAL INFORMATION:
/ APPLICANT: DOMNER, F.
/ APPLICANT: SCHEIFLINGER, F.
/ APPLICANT: FALKNER, F. G.
/ TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
/ NUMBER OF SEQUENCES: 52
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Foley & Lardner
/ STREET: 1800 Diagonal Road, Suite 500
/ CITY: Alexandria
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22313-0299
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/232,463
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/935,313
/ FILING DATE:
/ APPLICATION NUMBER: EP 91 114 300.6
/ FILING DATE: 26-AUG-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: 30472/114 IMMUG
/ TELECOMMUNICATION INFORMATION:
```

TELEPHONE: (703)636-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-F18
US-08-232-463-14

Query Match 1.8%; Score 43.8; DB 2; Length 7218;
Best Local Similarity 8.4%; Pred. No. 0.49; Mismatches 191; Indels 0; Gaps 0;
Matches 36; Conservative 204; Mismatches 191; Indels 0; Gaps 0;

Qy 628 TATCTGTGTAAAGCAATTCTTTCAGCATTCAGATTACATTTTGAAGCTAATA 687
Db 1474 TATCTATGACAGTAGTAAAGATAGAAATTTGTACRRRRRRRRRRRRRRRR 1415
Qy 688 GACGACGATTTGGTCCGTCATAGAAAGACTAATATATCTGAGTTAGTACA 747
Db 1414 RRR 1355
Qy 748 GCAGATTAGCACAGCAAAATTTGCTCAGTTTCAAGTAGCACTATCTGGGGAAGA 807
Db 1354 RRR 1295
Qy 808 GAGCTGAGCCAGTGTGCTCATTTTCTGATTAATCTTCAACATTTAAACCTGGATC 867
Db 1294 RRR 1235
Qy 868 TATGGAATCAAACCTGGGTAATAATCTTACGACATCACTACTGTAGGAAT 927
Db 1234 RRR 1175
Qy 928 GGACAGAAACAGACATTCAGTAATGGCTATATATAGAAATACGTAGAGGTGCC 987
Db 1174 RRR 1115
Qy 988 TGAATTAGACTTATTAAGAGTAGAGACAGAAATGAAATATCATGCAATTTCT 1047
Db 1114 RRR 1055
Qy 1048 GTAGCTCAGCA 1058
Db 1054 CTCGACCTGCA 1044

RESULT 10
US-09-621-976-2813
Sequence 2813, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.ppm
SEQ ID NO 2813
LENGTH: 832
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 235..399
US-09-621-976-2813

Query Match 1.8%; Score 43; DB 3; Length 832;
Best Local Similarity 15.0%; Pred. No. 0.31;
Matches 51; Conservative 147; Mismatches 142; Indels 1; Gaps 1;

Qy 1073 TTCGAAACGACCGAGTTTCCCAAACTACTGTGATGTTTCAGTGAATCTTCATCTC 1132
Db 35 KTYWRMRKKKKAAWKKWTWYRMWYMGYTKKAKCRKTKKKKKKKYMMWYGV 94
Qy 1133 ATGCTTATATGAGTAGATGATGATCTCAACCAAAATTAATGACAAAGACAGAT 1192
Db 95 RSTAAAMTRITWTGATYATRSMMYRRCMKKAYRRTTTCSSSGWMMGRKKAATW 154
Qy 1193 TGTGTTT-ATCTGTGGTAATACGTTTCTCAGTGTATAAAGACCTCCACGAG 1251
Db 155 WKKXTYVAATRYMMWMCWTRWPSWYCWMMGARSKSTWRSRSYASARSARCCYS 214
Qy 1252 TATAAGCTCTATGCAAGAAAGAAATGCAATATCTCTTACTCTCATTTATTTT 1311
Db 215 CSMGMSWKMWRMRMGATGAGMKAMRASCMRRRYAGKSTYSMMWCMTRSMKY 274
Qy 1312 CATTAGATAGCCGGTTTTTCTACTCAACTCAATTAAGATGAACAGATGAGGTAGT 1371
Db 275 CTTAKRTGYCYKKGWGRGRWYASKTKMKMKWCMARMTYRSTGTASMMWRTMY 334
Qy 1372 GACTGTTTATAAGAGATATAAGATTAATCATCAT 1412
Db 335 YTMWMMKWMKAMADBAWRMMWMMWAMRRACAAATATATATT 375

RESULT 11
US-09-949-016-18019/c
Sequence 18019, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C0001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 18019
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-18019

Query Match 1.8%; Score 42.6; DB 3; Length 601;
Best Local Similarity 54.0%; Pred. No. 0.34; Mismatches 74; Indels 0; Gaps 0;
Matches 87; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 2218 AGCTCCCTCACTACAGAAAAGTCACATTAATAATGCAACATGATGTTCTATTTGTT 2277
Db 258 AGTTTCTTCCACCTCAAAATTAACAAACAAACAACTTACTTGAAAAATTAACACTTC 199
Qy 2278 TTCTCTGCTGAGTGAATGATTAATATATTTTGTGGGGTGAATGATGA 2337
Db 198 CTATGAGATTGACTTATATTTCTCAATGCTTACCTTTACAGGTATTAATADTGA 139
Qy 2338 AAAAGAAATTAGAGCCACAGATTAATCTGAAGCTCAGATTT 2378
Db 138 AAGGAAGCTTGACGCTCATGACATTTGAAGCTGACATTT 98

RESULT 12
US-09-949-016-18020/c

```
; Sequence 18020, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 18020
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-18020
```

```
Query Match
Best Local Similarity 1.8%; Score 42.6; DB 3; Length 601;
Matches 87; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
```

```
QY 2218 AGCTCCCTCAACTACAGAAAAGTCACATTAATAAATGCAACATGATGTTCTATTGTT 2277
DB 554 AGTTTCTTCCACCTCCAAATTAACAAACAAACATACCTGGAAAAATTAACACTTC 495
QY 2278 TTCTCTGCTGATGATTAATGATTAATTAATTTTATTTTATGCGGTGAATGTGATGA 2337
DB 494 CTATGGATTTGACTTTATTTTCTCATTTGCTTCTTACAGGTGTTAATATGATGA 435
QY 2338 AAAAGAAAGTTAGAGCCCAAGATTAATCTGAAGCTCACTATT 2378
DB 434 AAAAGAAAGTTGACGCTCATGACATTTGAAGCTGACAAATT 394
```

```
RESULT 13
US-09-949-016-161279/c
; Sequence 161279, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 161279
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-161279
```

```
Query Match
Best Local Similarity 1.8%; Score 42.6; DB 3; Length 601;
Matches 87; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
```

```
QY 2218 AGCTCCCTCAACTACAGAAAAGTCACATTAATAAATGCAACATGATGTTCTATTGTT 2277
DB 258 AGTTTCTTCCACCTCCAAATTAACAAACAAACATACCTGGAAAAATTAACACTTC 199
```

```
QY 2278 TTCTCTGCTGATGATTAATGATTAATTAATTTTATTTTATGCGGTGAATGTGATGA 2337
DB 198 CTATGGATTTGACTTTATTTTCTCATTTGCTTCTTACAGGTGTTAATATGATGA 139
QY 2338 AAAAGAAAGTTAGAGCCCAAGATTAATCTGAAGCTCACTATT 2378
DB 138 AAAAGAAAGTTGACGCTCATGACATTTGAAGCTGACAAATT 98
```

```
RESULT 14
US-09-949-016-161280/c
; Sequence 161280, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 161280
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-161280
```

```
Query Match
Best Local Similarity 1.8%; Score 42.6; DB 3; Length 601;
Matches 87; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
```

```
QY 2218 AGCTCCCTCAACTACAGAAAAGTCACATTAATAAATGCAACATGATGTTCTATTGTT 2277
DB 554 AGTTTCTTCCACCTCCAAATTAACAAACAAACATACCTGGAAAAATTAACACTTC 495
QY 2278 TTCTCTGCTGATGATTAATGATTAATTAATTTTATTTTATGCGGTGAATGTGATGA 2337
DB 494 CTATGGATTTGACTTTATTTTCTCATTTGCTTCTTACAGGTGTTAATATGATGA 435
QY 2338 AAAAGAAAGTTAGAGCCCAAGATTAATCTGAAGCTCACTATT 2378
DB 434 AAAAGAAAGTTGACGCTCATGACATTTGAAGCTGACAAATT 394
```

```
RESULT 15
US-09-385-982-376/c
; Sequence 376, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCDNA-260XX
; CURRENT FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: 60/328,111
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 60/117,393
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: 60/098,639
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 376
; LENGTH: 611
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) ..(611)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-376

```

```

Query Match      1.8%; Score 42.6; DB 3; Length 611;
Best Local Similarity 55.0%; Pred. No. 0.35;
Matches 72; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

```

```

QY 2220 CTCCTCACTACAGAAAAAGTCACATRAAAATGCAACATGATGTTCTATTTGTTT 2279
    |||||
Db 351 CNCNTTATATTCNGTTTACGAAACNAAAAATGCGNACNTTTTTTTTTTTTTT 292
    |||||
QY 2280 TCTCTGCTTGATGTTAATTGATTATTATTATTTTATGCGCGTGATGTGATGAA 2339
    |||||
Db 291 TTTTNTTGTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTATAGATTATTATTAATA 232
    |||||
QY 2340 AGAAAGTTAGG 2350
    |||||
Db 231 GGGGAGATAGG 221

```

Search completed: May 27, 2006, 19:25:36
 Job time : 477.706 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2006, 18:56:27 ; Search time 3252.58 Seconds
(without alignments)
8994.972 Million cell updates/sec

Title: US-10-099-663-1
Perfect score: 2381
Sequence: 1 agcttcctgcgcagagaag.....atctgaagctcattcag 2381

Scoring table: IDENTITY NUC
Gapop 10.0, Gapept 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA_Main:

1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2381	100.0	2381	7 US-10-099-663-1	Sequence 1, Appli
2	336	14.1	336	7 US-10-099-663-2	Sequence 2, Appli
3	50.6	2.1	16914	8 US-10-741-601-5698	Sequence 2, Appli
4	50.6	2.1	16914	9 US-10-741-600-17777	Sequence 17777, A
5	50.6	2.1	16914	10 US-10-995-561-13349	Sequence 13349, A
6	50	2.1	564	16 US-11-136-527-3784	Sequence 3784, Ap
7	50	2.1	564	16 US-11-136-527-7880	Sequence 7880, Ap
8	49.8	2.1	630	6 US-10-027-632-5694	Sequence 5694, Ap
9	49.8	2.1	630	7 US-10-027-632-5694	Sequence 5694, Ap
10	49.8	2.1	5204	7 US-10-116-275-3533	Sequence 307, App
11	49.2	2.1	426	13 US-11-085-792-3	Sequence 3, Appli
12	48	2.0	546	4 US-09-925-065A-51803	Sequence 51803, A
13	48	2.0	546	5 US-09-925-065A-51803	Sequence 51803, A
14	48	2.0	546	12 US-10-301-480-153041	Sequence 153041, A
15	48	2.0	546	12 US-10-301-480-76450	Sequence 76450, A
16	48	2.0	568	12 US-10-301-480-288831	Sequence 288831, A
17	48	2.0	568	12 US-10-301-480-902240	Sequence 902240, A

C 18	48	2.0	569	4 US-09-925-065A-200196	Sequence 200196, A
C 19	48	2.0	569	5 US-09-925-065A-200196	Sequence 200196, A
C 20	48	2.0	570	4 US-09-925-065A-200197	Sequence 200197, A
C 21	48	2.0	570	5 US-09-925-065A-200197	Sequence 200197, A
C 22	48	2.0	599	12 US-10-301-480-288830	Sequence 288830, A
C 23	48	2.0	599	12 US-10-301-480-302239	Sequence 902239, A
C 24	48	2.0	814	4 US-09-925-065A-63272	Sequence 63272, A
C 25	48	2.0	814	5 US-09-925-065A-63272	Sequence 63272, A
C 26	48	2.0	814	12 US-10-301-480-164510	Sequence 164510, A
C 27	48	2.0	814	12 US-10-301-480-779319	Sequence 779319, A
C 28	48	2.0	991	12 US-10-301-480-584425	Sequence 584425, A
C 29	48	2.0	991	12 US-10-301-480-1197834	Sequence 1197834, A
C 30	48	2.0	992	12 US-10-301-480-584426	Sequence 584426, A
C 31	48	2.0	992	12 US-10-301-480-1197835	Sequence 1197835, A
C 32	47.6	2.0	399	13 US-11-085-792-1	Sequence 1, Appli
C 33	47.2	2.0	337	3 US-09-960-352-6036	Sequence 6036, Ap
C 34	47	2.0	2636	13 US-11-097-143-21061	Sequence 21061, A
C 35	46.8	2.0	662	9 US-10-335-053-36	Sequence 36, Appli
C 36	46.8	2.0	670	10 US-10-764-420-2245	Sequence 2245, Ap
C 37	46.2	1.9	6222	7 US-10-311-455-665	Sequence 665, App
C 38	44.4	1.9	565	4 US-09-925-065A-110075	Sequence 310075, A
C 39	44.4	1.9	565	5 US-09-925-065A-110075	Sequence 310075, A
C 40	44.4	1.9	573	12 US-10-301-480-385620	Sequence 385620, A
C 41	44.4	1.9	573	12 US-10-301-480-999029	Sequence 999029, A
C 42	44.2	1.9	515	3 US-09-871-161-474	Sequence 474, App
C 43	44.2	1.9	2252	8 US-10-741-601-261	Sequence 261, App
C 44	44.2	1.9	2252	9 US-10-741-601-745	Sequence 745, App
C 45	44.2	1.9	2252	10 US-10-995-561-449	Sequence 449, App

ALIGNMENTS

RESULT 1	US-10-099-663-1
Sequence 1, Application US/10099663	
Publication No. US20030177516m1	
GENERAL INFORMATION:	
APPLICANT: Avigenics, Inc	
TITLE OF INVENTION: Avian Gut-Specific Promoters	
FILE REFERENCE: A181	
CURRENT APPLICATION NUMBER: US/10/099, 663	
CURRENT FILING DATE: 2002-03-14	
NUMBER OF SEQ ID NOS: 19	
SOFTWARE: PatentIn version 3.0	
SEQ ID NO 1	
LENGTH: 2381	
TYPE: DNA	
ORGANISM: Gallus gallus	
FEATURE:	
NAME/KEY: 5'UTR	
LOCATION: (1)..(1626)	
NAME/KEY: exon	
LOCATION: (1627)..(1693)	
NAME/KEY: Intron	
LOCATION: (1694)..(2322)	
NAME/KEY: exon	
LOCATION: (2323)..(2381)	
Query Match	100.0%; Score 2381; DB 7; Length 2381;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 AGTTCTGCGCAGAAAAGCTGGGGTTTGTTCCTCAGACGTTAAGCAATC 60
DB	1 AGTTCTGCGCAGAAAAGCTGGGGTTTGTTCCTCAGACGTTAAGCAATC 60
QY	CCCAAGTCAAAAGCGCTGTAAAGAGAGATGCTCACTTCAATGAATGAATTG 120
DB	CCCAAGTCAAAAGCGCTGTAAAGAGAGATGCTCACTTCAATGAATGAATTG 120
QY	AAATATCATAAACGAGCTCTGTGGCAGATCAGAGTAACCTCTGCTGGACAAAT 180

```
Db 121 AAATATATATAAAAGAGCTCTGTGGAGATGAGATTAACCTCTCTGGAGCAAAATTT 180
Qy 181 CTTAAAGGTATAGAGTGAACAAGAGGTTTGGCACTAAATGAGCTAGATTGGACACA 240
Db 181 CTTAAAGGTATAGAGTGAACAAGAGGTTTGGCACTAAATGAGCTAGATTGGACACA 240
Qy 241 TTGATCTTCTAGAGAGCAAAAGGCTGGAAGAAACAATTAAATCTGGTGCACAGTCACT 300
Db 241 TTGATCTTCTAGAGAGCAAAAGGCTGGAAGAAACAATTAAATCTGGTGCACAGTCACT 300
Qy 301 AGCAGCCTGTTTGGGTGCACTACAGCACTTTGTTGGCAACATTAACAATCTAAGTTG 360
Db 301 AGCAGCCTGTTTGGGTGCACTACAGCACTTTGTTGGCAACATTAACAATCTAAGTTG 360
Qy 361 TTTTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Db 361 TTTTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Qy 421 TACTTCCCTCTGCACTCTCTTACCCAGATTAGCATTGATTTCAAAATGAACCTGAGTGA 480
Db 421 TACTTCCCTCTGCACTCTCTTACCCAGATTAGCATTGATTTCAAAATGAACCTGAGTGA 480
Qy 481 ATGGAAGGCCACCTAATTTTGTATGACAGCAAAAGTTCTAAATTTATAGTTATCTTCA 540
Db 481 ATGGAAGGCCACCTAATTTTGTATGACAGCAAAAGTTCTAAATTTATAGTTATCTTCA 540
Qy 541 GTAAAACTTTTGTCTGAGGTCTGGAAGAAAAAGAAATTTATGATTAACCCAGCTAGTA 600
Db 541 GTAAAACTTTTGTCTGAGGTCTGGAAGAAAAAGAAATTTATGATTAACCCAGCTAGTA 600
Qy 601 AAATCATTAAGTGAAGCCACCTGTTATCTGTGTGATAGCAACATTCATTTGAGCAT 660
Db 601 AAATCATTAAGTGAAGCCACCTGTTATCTGTGTGATAGCAACATTCATTTGAGCAT 660
Qy 661 TCAGAGTTTACATTTTGAAGCTAATAGACAGCATTTGGTCCGTCATAGGAACGA 720
Db 661 TCAGAGTTTACATTTTGAAGCTAATAGACAGCATTTGGTCCGTCATAGGAACGA 720
Qy 721 CTAACTAATATCTGAGTTTATGATCAAGCAATTTAGCAGCAACAAATTTGCTCAGTTTC 780
Db 721 CTAACTAATATCTGAGTTTATGATCAAGCAATTTAGCAGCAACAAATTTGCTCAGTTTC 780
Qy 781 AAGTGAACCTATCTGTGGGAAAGAGAGGCTGAGCCAGTGTGTGCTCATTTTCTGCAAT 840
Db 781 AAGTGAACCTATCTGTGGGAAAGAGAGGCTGAGCCAGTGTGTGCTCATTTTCTGCAAT 840
Qy 841 ATCTTCAACATTTTAAACCTGGGATCTATGGAATCAAAACGTTGGGTAAATTCAT 900
Db 841 ATCTTCAACATTTTAAACCTGGGATCTATGGAATCAAAACGTTGGGTAAATTCAT 900
Qy 901 TAGCAGCATCACTACTGTAGGAATGAGCAAGAACAGACATTCCTGAATGGGCTAT 960
Db 901 TAGCAGCATCACTACTGTAGGAATGAGCAAGAACAGACATTCCTGAATGGGCTAT 960
Qy 961 AAATTAAGAAATAGAGAGGTCTCTGAATTTAGACTACTAATTAAGAGTGAAGCA 1020
Db 961 AAATTAAGAAATAGAGAGGTCTCTGAATTTAGACTACTAATTAAGAGTGAAGCA 1020
Qy 1021 CGAATGAGAAATATCATGCAATTTCTGTAGCTCAGCACTAGACTGAAGGTTCTGA 1080
Db 1021 CGAATGAGAAATATCATGCAATTTCTGTAGCTCAGCACTAGACTGAAGGTTCTGA 1080
Qy 1081 CTGAACCGAGTTTCCCAAATCACTGTGATGTTCACTGAGATCTTCAATCTAGTTAT 1140
Db 1081 CTGAACCGAGTTTCCCAAATCACTGTGATGTTCACTGAGATCTTCAATCTAGTTAT 1140
Qy 1141 TAGTGAAGTGAATAGATTCTCAACAATTAAGATGAACAAGCAGAGATTGTGTTT 1200
Db 1141 TAGTGAAGTGAATAGATTCTCAACAATTAAGATGAACAAGCAGAGATTGTGTTT 1200
Qy 1201 ATCTGTGGTAAATAGTTTCTCAAGTTGTAATGAACCTCCCAACAGATTAAGTC 1260
Db 1201 ATCTGTGGTAAATAGTTTCTCAAGTTGTAATGAACCTCCCAACAGATTAAGTC 1260
Db 1201 ATCTGTGGTAAATAGTTTCTCAAGTTGTAATGAACCTCCCAACAGATTAAGTC 1260
Qy 1261 CTATGCAACAAGAAATATGCAATTAATCTCTTAAAGTCTCATTTATTTTCAATAGATA 1320
Db 1261 CTATGCAACAAGAAATATGCAATTAATCTCTTAAAGTCTCATTTATTTTCAATAGATA 1320
Qy 1321 GCCGGTTTTTATCAACAATCAATTAAGATGAACAAGATAGGTTAGTGACTGTTT 1380
Db 1321 GCCGGTTTTTATCAACAATCAATTAAGATGAACAAGATAGGTTAGTGACTGTTT 1380
Qy 1381 TAAAGAAAGTAAATTAAGATTAATCTATCTATTTGAGCAATTAAGAGGAGATTCAG 1440
Db 1381 TAAAGAAAGTAAATTAAGATTAATCTATCTATTTGAGCAATTAAGAGGAGATTCAG 1440
Qy 1441 CAAACAGTGTCTTACAAAGTGAAGAAACAAGTTAACTAAAGTGAACCCCTCTGACAA 1500
Db 1441 CAAACAGTGTCTTACAAAGTGAAGAAACAAGTTAACTAAAGTGAACCCCTCTGACAA 1500
Qy 1501 GATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATATGTAATTTGCTTGATA 1560
Db 1501 GATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATATGTAATTTGCTTGATA 1560
Qy 1561 AGCCTGTTCAATAATCTCTTGGCAAGCTCTGCTAATCAAGAAAGTCTGCTTACAG 1620
Db 1561 AGCCTGTTCAATAATCTCTTGGCAAGCTCTGCTAATCAAGAAAGTCTGCTTACAG 1620
Qy 1621 AGAAAGTGCATTTAAGGGTATGGAAGAAATGAAGAACTAATGAAGAAATTC 1680
Db 1621 AGAAAGTGCATTTAAGGGTATGGAAGAAATGAAGAACTAATGAAGAAATTC 1680
Qy 1681 ATGAAGCAATGGGTGAAGCTTACTTTTGAATGCTTCTTAAAGCAGATACCACTAC 1740
Db 1681 ATGAAGCAATGGGTGAAGCTTACTTTTGAATGCTTCTTAAAGCAGATACCACTAC 1740
Qy 1741 GCGGGAATCAAACTTAAGCTGTATGAACTAATCAATCAATCTGCTCTTGTGT 1800
Db 1741 GCGGGAATCAAACTTAAGCTGTATGAACTAATCAATCAATCTGCTCTTGTGT 1800
Qy 1801 CTGCTATTTTGGCTTGGCAATTTGCTTGCACCTTATTTGAAAGACTCTATAGAGGGA 1860
Db 1801 CTGCTATTTTGGCTTGGCAATTTGCTTGCACCTTATTTGAAAGACTCTATAGAGGGA 1860
Qy 1861 ATACAAGAAAGAAACATTCGATTTTATTTGCAATGCGAATATCTTATAGCATTTAGCT 1920
Db 1861 ATACAAGAAAGAAACATTCGATTTTATTTGCAATGCGAATATCTTATAGCATTTAGCT 1920
Qy 1921 AATTCAGTAAAGGATTCGAGCAAAATTTAAATGAATTAATGTAGAAATTAAT 1980
Db 1921 AATTCAGTAAAGGATTCGAGCAAAATTTAAATGAATTAATGTAGAAATTAAT 1980
Qy 1981 TTGATTAAGACTGTTGAAATAATTAACAAGAGGGAATTTGCTGCTCCAGTTTGCAG 2040
Db 1981 TTGATTAAGACTGTTGAAATAATTAACAAGAGGGAATTTGCTGCTCCAGTTTGCAG 2040
Qy 2041 ACAACATGATTTAGTCAATTTTAAACATGCTAGTCTTACTTTAAGCTTGAACAATC 2100
Db 2041 ACAACATGATTTAGTCAATTTTAAACATGCTAGTCTTACTTTAAGCTTGAACAATC 2100
Qy 2101 CTGTAATATGAGTAAACATTAACATCTTCTAGTTGATAGTATGATTTTCAAGCTGA 2160
Db 2101 CTGTAATATGAGTAAACATTAACATCTTCTAGTTGATAGTATGATTTTCAAGCTGA 2160
Qy 2161 ACACCTGCTCAGTGAAGGTGAGAAAGATGAACCTGAGTCAAGATTTCTGGGCTAAGC 2220
Db 2161 ACACCTGCTCAGTGAAGGTGAGAAAGATGAACCTGAGTCAAGATTTCTGGGCTAAGC 2220
Qy 2221 TCCCTCAACTACAGAAAGATCAATTAATAATGAACATGATGTTCTAATTTGTTT 2280
Db 2221 TCCCTCAACTACAGAAAGATCAATTAATAATGAACATGATGTTCTAATTTGTTT 2280
Qy 2281 CTCTGCTGATGTTAATGATTAATTAATTTTCTTAAAGCTGGAATGAGTGAAGAA 2340
Db 2281 CTCTGCTGATGTTAATGATTAATTAATTTTCTTAAAGCTGGAATGAGTGAAGAA 2340
Db 2281 CTCTGCTGATGTTAATGATTAATTAATTTTCTTAAAGCTGGAATGAGTGAAGAA 2340
```

OY 2341 GAAAGTTAGAGCCGACGATTAATCTGAAGCTCACTATTCAAG 2381
 Db 2241 GAAAGTTAGAGAGCCGACGATTAATCTGAAGCTCACTATTCAAG 2281

RESULT 2

US-10-099-663-2
 ; Sequence 2, Application US/10099663
 ; Publication No. US20030177516A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Avigenics, Inc
 ; TITLE OF INVENTION: Avian Gut-Specific Promoters
 ; FILE REFERENCE: A181
 ; CURRENT APPLICATION NUMBER: US/10/099,663
 ; CURRENT FILING DATE: 2002-03-14
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 336
 ; TYPE: DNA
 ; ORGANISM: Gallus gallus
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(336)
 ; OTHER INFORMATION: chick intestinal fatty acid binding protein promoter region
 US-10-099-663-2

Query Match 14.1%; Score 336; DB 7; Length 336;
 Best Local Similarity 100.0%; Pred. No. 8.6e-74;
 Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1301 ATTATATTTTTCATTAGATAGCCGGTTTTTCTTACTACAACTCAATTAAGATGACAGAAATG 1360
 Db 1 ATTATATTTTTCATTAGATAGCCGGTTTTTCTTACTACAACTCAATTAAGATGACAGAAATG 60
 OY 1361 AATGGTTAGTACTGTTTATTAAGAGATTAATTAAGATTAATCATCATTTTGAAGCAA 1420
 Db 61 AATGGTTAGTACTGTTTATTAAGAGATTAATTAAGATTAATCATCATTTTGAAGCAA 120
 OY 1421 TAAGGGAGGAGAGATTCAGCAACAGTGTCTTCAAGTGTGAAAACAAAGTTAAACTTAA 1480
 Db 121 TAAGGGAGGAGAGATTCAGCAACAGTGTCTTCAAGTGTGAAAACAAAGTTAAACTTAA 180
 OY 1481 GTGACCCCTCTCTTGAAGATCAATGCAAGTGTGAGCTTTAGCCAGCAATCATCA 1540
 Db 181 GTGACCCCTCTCTTGAAGATCAATGCAAGTGTGAGCTTTAGCCAGCAATCATCA 240
 OY 1541 TGTAAATGCTTTCCTGATTAAGCCTGTTTCAATTAATCTTTGGAAAGCTCTGCTACTTA 1600
 Db 241 TGTAAATGCTTTCCTGATTAAGCCTGTTTCAATTAATCTTTGGAAAGCTCTGCTACTTA 300
 OY 1601 CCAGAACTGTGCTTACAGACAGAAAGATGCAATTA 1636
 Db 301 CCAGAACTGTGCTTACAGACAGAAAGATGCAATTA 336

RESULT 3

US-10-741-601-5698
 ; Sequence 5698, Application US/10741601
 ; Publication No. US20040166519A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: C1001500
 ; CURRENT APPLICATION NUMBER: US/10/741,601
 ; CURRENT FILING DATE: 2003-12-22
 ; NUMBER OF SEQ ID NOS: 26415
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5698
 ; LENGTH: 16914
 ; TYPE: DNA

; ORGANISM: Homo sapiens
 US-10-741-601-5698

Query Match 2.1%; Score 50.6; DB 8; Length 16914;
 Best Local Similarity 58.8%; Pred. No. 0.69;
 Matches 124; Conservative 2; Mismatches 76; Indels 9; Gaps 2;

OY 1516 TGACCTTTAGCCAGCAGATCATGATTAATTTCTTCTGATTAAGCTTTCATTAAT 1575
 Db 5945 TGAATCTTAAGCTTCCACATCAGATTAATGATTTGATTAAGATTAATTAATTAAT 6004
 OY 1576 TCTC-----TTTGCAAGCTCTGCTACTTACCAAGTCTGCTTACAGAGAA 1627
 Db 6005 TCTGCCCCAAGAGAGAGCTGAATCTTCTAGCTGCTTGAAGGCTACT-CAACTGAAATCA 6063
 OY 1628 TGGCAATTTAAGGTAAGTGAAGAAATAGAGAAATAGAACTATGAAATTTCAATGAA 1687
 Db 6064 TGGCGTTTGACAGACTTGGAGAGTACCGAGTGAAGAACTTAATGACAGTTTCAATGAAA 6123
 OY 1688 CAATGGGTAAGCTTACTTTTGTGAATGCTT 1718
 Db 6124 AAATGGGTAAGACTTTATTTCTTTGTGGCT 6154

RESULT 4

US-10-741-600-17777
 ; Sequence 17777, Application US/10741600
 ; Publication No. US20050026169A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: C1001499
 ; CURRENT APPLICATION NUMBER: US/10/741,600
 ; CURRENT FILING DATE: 2003-12-22
 ; NUMBER OF SEQ ID NOS: 7397
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17777
 ; LENGTH: 16914
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-741-600-17777

Query Match 2.1%; Score 50.6; DB 9; Length 16914;
 Best Local Similarity 58.8%; Pred. No. 0.69;
 Matches 124; Conservative 2; Mismatches 76; Indels 9; Gaps 2;

OY 1516 TGACCTTTAGCCAGCAGATCATGATTAATTTCTTCTGATTAAGCTTTCATTAAT 1575
 Db 5945 TGAATCTTAAGCTTCCACATCAGATTAATGATTTGATTAAGATTAATTAATTAAT 6004
 OY 1576 TCTC-----TTTGCAAGCTCTGCTACTTACCAAGTCTGCTTACAGAGAA 1627
 Db 6005 TCTGCCCCAAGAGAGAGCTGAATCTTCTAGCTGCTTGAAGGCTACT-CAACTGAAATCA 6063
 OY 1628 TGGCAATTTAAGGTAAGTGAAGAAATAGAGAAATAGAACTATGAAATTTCAATGAA 1687
 Db 6064 TGGCGTTTGACAGACTTGGAGAGTACCGAGTGAAGAACTTAATGACAGTTTCAATGAAA 6123
 OY 1688 CAATGGGTAAGCTTACTTTTGTGAATGCTT 1718
 Db 6124 AAATGGGTAAGACTTTATTTCTTTGTGGCT 6154

RESULT 5

US-10-995-561-13349
 ; Sequence 13349, Application US/10995561
 ; Publication No. US20050272054A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
 ; TITLE OF INVENTION: DETECTION AND USES THEREOF

FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13349
LENGTH: 16914
TYPE: DNA
ORGANISM: Homo sapiens
US-10-995-561-13349

Query Match 2.1%; Score 50.6; DB 10; Length 16914;
Best Local Similarity 58.8%; Pred. No. 0.69;
Matches 124; Conservative 2; Mismatches 76; Indels 9; Gaps 2;

QY 1516 TGAGCTTTAGCCAGCCATCATCATTAATTCCTTCGATAGACCTGTTCAATAAT 1575
DB 5945 TGAACCTTAAGCTTCACATCACTAGAACTTGTTAAGATTAAGAAATATTAAT 6004
QY 1576 TCTC-----TTGCAAGCTCTGCTACTTACAGAAAGTCTGCTACAGACAGAAAGA 1627
DB 6005 TCTGCCCCAAGACAGACCTGAATCTAGCTGCTAGAGGCTGACT-CAACTGAAATCA 6063
QY 1628 TGGCATTTAAGCGTACTTGAATAATGAGAAAAATGAGAACTATGAAAAATTCATGGAAG 1687
DB 6064 TGGCGTTTGAAGCAGCACTTGAAGGTAAGCCGAGTGAATACTATGACAAAGTTCAATGGAAA 6123
QY 1688 CAATGGCTAAGCCTTACTTTTGAATGCT 1718
DB 6124 AATGGTAAGACTTTATTTCTTTTGCT 6154

RESULT 6

US-11-136-527-3784
Sequence 3784, Application US/11136527
Publication No. US20050287570A1
GENERAL INFORMATION:
APPLICANT: Wyeth
APPLICANT: Mounts, William M
TITLE OF INVENTION: Probe Arrays for Expression Profiling of Rat Genes
FILE REFERENCE: 031896-041000 (AM101086)
CURRENT APPLICATION NUMBER: US/11/136,527
CURRENT FILING DATE: 2005-05-25
PRIOR APPLICATION NUMBER: US 60/574,294
PRIOR FILING DATE: 2005-05-26
NUMBER OF SEQ ID NOS: 362830
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3784
LENGTH: 564
TYPE: DNA
ORGANISM: Rattus norvegicus
US-11-136-527-3784

Query Match 2.1%; Score 50; DB 16; Length 564;
Best Local Similarity 75.6%; Pred. No. 0.15;
Matches 62; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1615 ACAGACAGAAAGTGCATTTAAGCTATCTTGAAAAATGAGAAAAATGAACTATGAA 1674
DB 1 ACAGCTGACATCATGCTATTTGACACTTGGAAAGTGAACCGAATGAGAACTATGAA 60
QY 1675 AATTCATGAGCAATGGGTA 1696
DB 61 AAGTTCATGAGAAATGGGCA 82

RESULT 7

US-11-136-527-7880
Sequence 7880, Application US/11136527
Publication No. US20050287570A1
GENERAL INFORMATION:
APPLICANT: Wyeth
APPLICANT: Mounts, William M

TITLE OF INVENTION: Probe Arrays for Expression Profiling of Rat Genes
FILE REFERENCE: 031896-041000 (AM101086)
CURRENT APPLICATION NUMBER: US/11/136,527
CURRENT FILING DATE: 2005-05-25
PRIOR APPLICATION NUMBER: US 60/574,294
PRIOR FILING DATE: 2005-05-26
NUMBER OF SEQ ID NOS: 362830
SOFTWARE: PatentIn version 3.2
SEQ ID NO 7880
LENGTH: 564
TYPE: DNA
ORGANISM: Rattus norvegicus
US-11-136-527-7880

Query Match 2.1%; Score 50; DB 16; Length 564;
Best Local Similarity 75.6%; Pred. No. 0.15;
Matches 62; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1615 ACAGACAGAAAGTGCATTTAAGCTATCTTGAAAAATGAGAAAAATGAGAACTATGAA 1674
DB 1 ACAGCTGACATCATGCTATTTGACACTTGGAAAGTGAACCGAATGAGAACTATGAA 60
QY 1675 AATTCATGAGCAATGGGTA 1696
DB 61 AAGTTCATGAGAAATGGGCA 82

RESULT 8

US-10-027-632-5694/C
Sequence 5694, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
POLYMORPHISMS IN THE HUMAN GENOME
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5694
LENGTH: 630
TYPE: DNA
ORGANISM: Human
US-10-027-632-5694

Query Match 2.1%; Score 49.8; DB 6; Length 630;
Best Local Similarity 59.2%; Pred. No. 0.18;
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

QY 1516 TGAGCTTGAAGCAGCAATCATCATTAATTCCTTCGATAGACCTGTTCAATAAT 1575
DB 531 TGAACCTTAAGCTTCACATCACTAGAACTTGTTCAAGATTAAGAAATATTAAT 472
QY 1576 TCTC-----TTGCAAGCTCTGCTACTTACAGAAAGTCTGCTACAGACAGAAAGA 1627
DB 471 TCTGCCCCAGACAGACCTGAATCTAGCTGCTAGAGGCTGACT-CAACTGAAATCA 413
QY 1628 TGGCATTTAAGCGTACTTGAATAATGAGAAAAATGAGAACTATGAAAAATTCATGGAAG 1687

```

Db      412 TGGCGTTGACAGCAGCTTGAGAGTAGACCGAGAGTGAAGAACTATGACAGTTTCATGAGAA 353
Qy      1688 CAATGGGTAAAGCCTTACTTTTGTGATGCT 1718
Db      352 AAATGGGTAAAGACTTATTTCTTTGTGCT 322

```

RESULT 9

```

US-10-027-632-5694/C
; Sequence 5694, Application US/10027632
; Publication No. US2003020407549
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5694
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-5694

```

Query Match 2.1%; Score 49.8; DB 7; Length 630;

Best Local Similarity 59.2%; Pred. No. 0.18;

Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

```

Qy      1516 TGAGCTTTAGCCAGCAGCATCATCATGTAATGCTTTCCTGATAGCCTGTTCTAAT 1575
Db      531 TGAACCTTAAAGCTTCCACATCAAGTATGAGTTGGTTCAAGATTAAGAAATATATTAAT 472
Qy      1576 TCTC-----TTGCAAGCTCTGCTACTTACCAAGAGTCTGCTACAGACAGAAAGA 1627
Db      471 TCTGCCCCAAGACAGACCTGAATCTCTAGCTGCTTAAAGGCTACT-CAACTGAAATCA 413
Qy      1628 TGGCATTTAACGGTACTTGAAAAATAGAAAAATGAGAACTATGAAAAATTCATGAG 1687
Db      412 TGGCGTTTGAACAGCACTTGAAAGTAGACCGAGTGAAGAAACTATGACAAAGTTTCATGAAA 353
Qy      1688 CAATGGGTAAAGCCTTACTTTTGTGATGCT 1718
Db      352 AAATGGGTAAAGACTTATTTCTTTGTGCT 322

```

RESULT 10

```

US-10-116-275-303
; Sequence 303, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Linda
; APPLICANT: Higgins, Lisa

```

```

; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 303
; LENGTH: 5204
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-116-275-303

```

Query Match 2.1%; Score 49.8; DB 7; Length 5204;

Best Local Similarity 59.2%; Pred. No. 0.57;

Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

```

Qy      1516 TGAGCTTTAGCCAGCAGCATCATCATGTAATGCTTTCCTGATAGCCTGTTCTAAT 1575
Db      971 TGAACCTTAAAGCTTCCACATCAAGTATGAGTTGGTTCAAGATTAAGAAATATATTAAT 1030
Qy      1576 TCTC-----TTGCAAGCTCTGCTACTTACCAAGAGTCTGCTACAGACAGAAAGA 1627
Db      1031 TCTGCCCCAAGACAGACCTGAATCTCTAGCTGCTTAAAGGCTACT-CAACTGAAATCA 1089
Qy      1628 TGGCATTTAACGGTACTTGAAAAATAGAAAAATGAGAACTATGAAAAATTCATGAG 1687
Db      1090 TGGCGTTTGAACAGCACTTGAAAGTAGACCGAGTGAAGAAACTATGACAAAGTTTCATGAAA 1149
Qy      1688 CAATGGGTAAAGCCTTACTTTTGTGATGCT 1718
Db      1150 AAATGGGTAAAGACTTATTTCTTTGTGCT 1180

```

RESULT 11

```

US-11-085-792-3
; Sequence 3, Application US/11085792
; Publication No. US20050244864A1
; GENERAL INFORMATION:
; APPLICANT: Kleinfeld, Alan
; TITLE OF INVENTION: FLUORESCENT PROBES FOR HYDROPHOBIC
; TITLE OF INVENTION: ANALYTES
; FILE REFERENCE: PFASC.065PR
; CURRENT APPLICATION NUMBER: US/11/085,792
; CURRENT FILING DATE: 2005-03-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 426
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (1)...(426)
; OTHER INFORMATION: rat intestinal fatty acid binding protein DNA
; OTHER INFORMATION: sequence coding for substitution of alanine for
; OTHER INFORMATION: leucine at position 72
; OTHER INFORMATION: 3' terminus codes for a his tag
US-11-085-792-3

```

Query Match 2.1%; Score 49.2; DB 13; Length 426;

Best Local Similarity 81.4%; Pred. No. 0.2;

Matches 57; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

Qy      1627 ATGGCATTTAACGGTACTTGAAAAATAGAAAAATGAGAACTATGAAAAATTCATGAA 1686
Db      1 ATGGCATTTGAATGACACTTGAAAAAGTAGACCGGAATGAGAACTATGAAAAAGTTTCATGAG 60
Qy      1687 GCATGGGTA 1696
Db      61 AAATGGGCA 70

```

```
RESULT 12
US-09-925-065A-51803/c
; Sequence 51803, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51803
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-51803

Query Match
Best Local Similarity 68.8%; Score 48; DB 4; Length 546;
Matches 66; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1623 AAAGTGCATTAAACGGTACTTGAAGAAATAGAGAAATGAACTATGAAATTCAT 1682
DB 415 AATCATGGCGTTTGACACACTTGAAGGTAGACCGGAGTAAACTATGACAACTTCAT 356
QY 1683 GGAAGCAATGGGTAAAGCCTTACTTTTGAATGCCT 1718
DB 355 GGAATAATGGGTAAAGACTTATTTCTTGTGGCT 320

RESULT 13
US-09-925-065A-51803/c
; Sequence 51803, Application US/09925065A
; Publication No. US2005028172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51803
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-51803

Query Match
Best Local Similarity 68.8%; Score 48; DB 5; Length 546;
Matches 66; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1623 AAAGTGCATTAAACGGTACTTGAAGAAATAGAGAAATGAACTATGAAATTCAT 1682
DB 415 AATCATGGCGTTTGACACACTTGAAGGTAGACCGGAGTAAACTATGACAACTTCAT 356
QY 1683 GGAAGCAATGGGTAAAGCCTTACTTTTGAATGCCT 1718
DB 355 GGAATAATGGGTAAAGACTTATTTCTTGTGGCT 320

RESULT 14
US-10-301-480-153041/c
; Sequence 153041, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 153041
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-153041

Query Match
Best Local Similarity 68.8%; Score 48; DB 12; Length 546;
Matches 66; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1623 AAAGTGCATTAAACGGTACTTGAAGAAATAGAGAAATGAACTATGAAATTCAT 1682
DB 415 AATCATGGCGTTTGACACACTTGAAGGTAGACCGGAGTAAACTATGACAACTTCAT 356
QY 1683 GGAAGCAATGGGTAAAGCCTTACTTTTGAATGCCT 1718
DB 355 GGAATAATGGGTAAAGACTTATTTCTTGTGGCT 320

RESULT 15
US-10-301-480-766450/c
; Sequence 766450, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 766450
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-766450

Query Match
Best Local Similarity 68.8%; Score 48; DB 12; Length 546;
Matches 66; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
```


Qy 1623 AAAGATGGCATTTAAGGTACTTGAAAAATAGAAAAATGAACTATGAAAAATTCAT 1682
Db 415 AATCATGGCGTTTGACAGCACTTGAGAGGTAGACCGAGTGAACCTATGACAGTTTCAT 356
Qy 1683 GGAAGCAATGGGTAAAGCCTTACTTTTGAATGCCT 1718
Db 355 GGAATAAATGGGTAAAGACTTATTTCTTGTGGCT 320

Search completed: May 27, 2006, 20:03:04
Job time : 3254.58 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using SW model

Run on: May 27, 2006, 18:57:23 ; Search time 42.7971 Seconds
(without alignments)
6280.280 Million cell updates/sec

Title: US-10-099-663-1
Perfect score: 2381
Sequence: 1 agctcctgcgcagcaaaag5.....atcgaagctcattccag 2381

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 242596 seqs, 56442199 residues

Total number of hits satisfying chosen parameters: 485192

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_NA_New:*
1: /EMC_Celerra_SIDS3/pcodaca/1/pubpna/US09_NEW_PUB.seq:*
2: /EMC_Celerra_SIDS3/pcodaca/1/pubpna/US06_NEW_PUB.seq:*
3: /EMC_Celerra_SIDS3/pcodaca/1/pubpna/US07_NEW_PUB.seq:*
4: /EMC_Celerra_SIDS3/pcodaca/1/pubpna/US08_NEW_PUB.seq:*
5: /EMC_Celerra_SIDS3/pcodaca/1/pubpna/PC1_NEW_PUB.seq:*
6: /EMC_Celerra_SIDS3/pcodaca/1/pubpna/US10_NEW_PUB.seq:*
7: /EMC_Celerra_SIDS3/pcodaca/1/pubpna/US11_NEW_PUB.seq:*
8: /EMC_Celerra_SIDS3/pcodaca/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49.8	2.1	5204	US-10-524-021-12	Sequence 12, App1
2	38	1.6	1907	US-10-953-349-33745	Sequence 33745, A
3	38	1.6	2244	US-11-217-529-79170	Sequence 79170, A
4	37	1.6	129	US-10-488-619-2835	Sequence 2835, Ap
5	36.6	1.5	394191	US-10-506-549-3	Sequence 3, App1
6	35.8	1.5	4209	US-11-217-529-4068	Sequence 4068, Ap
7	35.2	1.5	2286	US-11-217-529-2925	Sequence 2925, Ap
8	35.2	1.5	3323	US-10-473-173-40	Sequence 40, App1
9	35	1.5	1275	US-11-217-529-2163	Sequence 2163, Ap
10	35	1.5	2439	US-10-953-349-37222	Sequence 37222, A
11	34.6	1.5	111	US-10-488-619-1326	Sequence 1326, Ap
12	34.6	1.5	2565	US-11-315-766-19	Sequence 19, App1
13	34.6	1.5	4692	US-10-511-937-646	Sequence 646, App
14	34.4	1.4	624	US-10-953-349-6180	Sequence 6180, Ap
15	34	1.4	989	US-10-953-349-15742	Sequence 15742, A
16	34	1.4	3426	US-10-511-937-438	Sequence 438, App
17	34	1.4	3877	US-10-196-749-263	Sequence 263, App
18	34	1.4	3877	US-11-101-316-71	Sequence 71, App1
19	33.8	1.4	346	US-10-511-937-529	Sequence 529, App
20	33.8	1.4	560	US-10-953-349-37522	Sequence 37522, A
21	33.8	1.4	750	US-10-953-349-1467	Sequence 1467, Ap
22	33.8	1.4	1146	US-10-953-349-20608	Sequence 20608, A
23	33.8	1.4	1535	US-10-953-349-15855	Sequence 15855, A
24	33.8	1.4	1582	US-11-222-810-10	Sequence 10, App1
25	33.8	1.4	1582	US-11-222-810-12	Sequence 12, App1

C	26	33.8	1.4	1921	6	US-10-953-349-9882	Sequence 9882, Ap
C	27	33.8	1.4	1953	7	US-11-293-697-2048	Sequence 2048, Ap
C	28	33.8	1.4	2205	6	US-10-953-349-9921	Sequence 9921, Ap
C	29	33.8	1.4	3038	7	US-11-293-697-471	Sequence 471, App
C	30	33.8	1.4	4640	6	US-10-196-749-75	Sequence 75, App1
C	31	33.6	1.4	1638	7	US-11-217-529-79683	Sequence 79683, Ap
C	32	33.6	1.4	2320	6	US-10-953-349-8153	Sequence 8153, Ap
C	33	33.6	1.4	2424	7	US-11-217-529-853	Sequence 853, App
C	34	33.6	1.4	4017	6	US-11-326-265-10	Sequence 10, App1
C	35	33.4	1.4	485	7	US-10-196-749-185	Sequence 185, App
C	36	33.4	1.4	485	7	US-11-101-316-43	Sequence 43, App1
C	37	33.4	1.4	883	6	US-10-953-349-15521	Sequence 15521, A
C	38	33.4	1.4	1969	6	US-10-196-749-193	Sequence 193, App
C	39	33.4	1.4	4670	7	US-11-145-3070-23	Sequence 29, App1
C	40	33.2	1.4	644	6	US-10-196-749-213	Sequence 85, App
C	41	33.2	1.4	644	7	US-11-101-316-55	Sequence 55, App1
C	42	33.2	1.4	780	6	US-10-953-349-36666	Sequence 36666, A
C	43	33.2	1.4	1035	7	US-11-217-529-82312	Sequence 82312, A
C	44	33.2	1.4	1566	7	US-11-217-529-79711	Sequence 79711, A
C	45	33.2	1.4	2082	7	US-11-293-697-698	Sequence 698, App

ALIGNMENTS

```
RESULT 1
US-10-524-021-12
; Sequence 12, Application US/10524021
; Publication No. US20060099590A1
; GENERAL INFORMATION:
; APPLICANT: NAGOYA INDUSTRIAL SCIENCE RESEARCH INSTITUTE
; APPLICANT: Gifu INTERNATIONAL INSTITUTE OF BIOTECHNOLOGY
; APPLICANT: YAMADA, Yoshiji
; APPLICANT: KOKOTA, Mitsuhiro
; TITLE OF INVENTION: Method for diagnosing a risk of restenosis after percutaneous cor
; FILE REFERENCE: C0200501
; CURRENT APPLICATION NUMBER: US/10/524, 021
; PRIOR FILING DATE: 2005-02-09
; PRIOR APPLICATION NUMBER: JP P2002-233041
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 5204
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-524-021-12

Query Match      2.1% Score 49.8; DB 6; Length 5204;
Best Local Similarity 59.2%; Pred. No. 0.0038;
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

QY 1516 TGACCTTAAAGCAGCAGCATATCATGTAATTCCTTCGTATAGCTGTTATTAAT 1575
    |||||
DB 971 TGAAGTTAAAGCTTCAATCAGATGAAGTTGTTCAAGATAAATATATAAT 1030
    |||||
QY 1576 TCTC-----TTGCAAAAGCTTGTCTACTTACCAAGTTCCTTACAGCAAGAA 1627
    |||||
DB 1031 TCTGCGCCAGCAGCAGCATGTAATCTAGTGTCTAGAGGCTACT-CAACTGAAATCA 1089
    |||||
QY 1528 TGGCATTTAAGCGGTAATGGAATAATAGAAAATAGAACTATGAAAATTCATGAG 1687
    |||||
DB 1090 TGGGTTTGAAGCAGCAGCTTGAAGGTAGCCGAGTGAACCTATGACAAAGTTTGA 1149
    |||||
QY 1688 CAATGGGTAAGCCTTACTTATTTTGAATGCTT 1718
    |||||
DB 1150 AAATGGGTAAGACTTATTTCTTTGTGGCT 1180
    |||||

RESULT 2
US-10-953-349-33745
; Sequence 33745, Application US/10953349
```

```

: Publication No. US20060107345A1
:
: GENERAL INFORMATION:
: APPLICANT: ALEXANDROV, Nikolai et al.
: TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
: TITLE OF INVENTION: ENCODED THERBY
: FILE REFERENCE: 2750-1579PUS2
: CURRENT APPLICATION NUMBER: US/10/953,349
: CURRENT FILING DATE: 2004-09-30
: NUMBER OF SEQ ID NOS: 40252
: SOFTWARE: PatentIn version 3.3
: SEQ ID NO 33745
: LENGTH: 1907
:
: TYPE: DNA
:
: ORGANISM: Zea mays subsp. mays
: US-10-953-349-33745

```

Query Match	1.6%	Score 38;	DB 6;	Length 1907;
Best Local Similarity	62.8%	Pred. NO. 2;		
Matches 59;	Conservative 0;	Mismatches 35;	Indels 0;	Gaps 0;

QY		2227	AACCTCAGAAAAAAGCCAAATTTAAAAGCAACAATGATGTTCATTGGTTTTCTCGC	2286
Dδ		1108	AACCTGCAGAAAAAAGCAAAAGCACGCCGAAGAAGCTTTTTTTTTTTTTTCTT	1167
QY		2287	TTGAGGTBAATGATTAATAATTTTTTTTTT	2320
Dβ		1168	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	1201

RESULT 3
US-91-217-529-79170

Sequence 79170, Application US/11217529
Publication No. US20060099612A1

1 GENERAL INFORMATION:
 2 APPLICANT: SUNTORY LIMITED
 3 APPLICANT: NAKAO, YOSHIHIRO
 4 APPLICANT: NAKAMURA, NORIHISA
 5 APPLICANT: KODAMA, YUKIO
 6 APPLICANT: FUJIMURA, TOMOKO
 7 APPLICANT: ASHIKARI, TOSHIHIKO
 8 TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
 9 FILE REFERENCE: 5-38-285
 10 CURRENT APPLICATION NUMBER: US/11/217,529
 11 PRIORITY FILING DATE: 2005-09-02
 12 PRIOR APPLICATION NUMBER: US 10/932,182
 13 PRIORITY FILING DATE: 2004-09-02
 14 NUMBER OF SEQ ID NOS: 197023
 15 SOFTWARE: PatentIn version 3.3
 16 SEQ ID NO 79170
 17 LENGTH: 2244
 18 TYPE: DNA
 19 ORGANISM: *Saccharomyces pastorianus*
 20 US-11-217-529-79170

```
Query Match      1.6%; Score 38; DB 7; Length 2244;
Best Local Similarity 50.5%; Pred. No. 2.1;
Matches 92; Conservative 0; Mismatches 90; Indels 0
```

Oy 523 ATTATAGTATTACTTCAGTAAACCTTTGCTCAGGTGGAGAAGAAAAGAGTAT 582
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2002 AGCTATTCMAACGCTGTGAAAAAGATCATGTGTAAATATGCCAAGACAAGTGAT 2061

[illegible]

64	CAACATATATTCACATTCGGGATTACATTTTCAGCTATATACAGCAGATTTG	702
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93		
94		
95		
96		
97		
98		
99		
100		
101		
102		
103		
104		
105		
106		
107		
108		
109		
110		
111		
112		
113		
114		
115		
116		
117		
118		
119		
120		
121		
122		
123		
124		
125		
126		
127		
128		
129		
130		
131		
132		
133		
134		
135		
136		
137		
138		
139		
140		
141		
142		
143		
144		
145		
146		
147		
148		
149		
150		
151		
152		
153		
154		
155		
156		
157		
158		
159		
160		
161		
162		
163		
164		
165		
166		
167		
168		
169		
170		
171		
172		
173		
174		
175		
176		
177		
178		
179		
180		
181		
182		
183		
184		
185		
186		
187		
188		
189		
190		
191		
192		
193		
194		
195		
196		
197		
198		
199		
200		
201		
202		
203		
204		
205		

Db 2182 TC 2183

RESULT 4
US-10-488-619-2835/c
; Sequence 2835, Application US/10486619
; Publication No. US20060095578A1
; Copyright information

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

; CURRENT APPLICATION NUMBER: US/10/488,619
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 3040
; SOFTWARE: PatentIn version 3.1

```

```

; SEQ ID NO 2835
;
; LENGTH: 129
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-488-619-2835

```

Query Match	1.6%;	Score 37;	DB 6;	Length 129;
Best Local Similarity	58.7%;	Pred. No. 1.2;		
Matches 64;	Conservative 0;	Mismatches 45;	Indels 0;	Gaps 0;

QY 2264 TGTTCATTTTGGTTTTCTCGCTGATGTAATTGATATATTTTTTTTAGG 2323
 | | | | | | | | | | | | | | | | | | | | | |
Db 118 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 59

QY 2324 CGTGAATGTCATGMAAGAAAGTTAGGACCCACGATTACTGAGCTC 2372

Db 58 TTTTATTATTAGGATAAGCTGTTTAGGTAGCACCAGAGAAATTTGAAATTC 10

RESULT 5
US-10-506-549-3

```

1 // Sequence 3, Application US/10506549
2 // Publication No. US20060100417A1
3 // GENERAL INFORMATION:
4 // APPLICANT: APPLERA CORPORATION
5 // TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
6 // TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
7 // TITLE OF INVENTION: AND USES THEREOF
8 // FILE REFERENCE: CL001361-US
9 // CURRENT APPLICATION NUMBER: US/10/506,549
10 // CURRENT FILING DATE: 2004-09-03
11 // PRIOR APPLICATION NUMBER: 60/361,343
12 // PRIOR FILING DATE: 2002-03-05
13 // NUMBER OF SEQ ID NOS: 4
14 // SOFTWARE: FastSeq for Windows Version 4.0
15 // SEQ ID NO 3
16 // LENGTH: 394191
17 // TYPE: DNA
18 // ORGANISM: Homo sapiens
19 // FEATURE:
20 // NAME/KEY: misc_feature
21 // LOCATION: (1)...(394191)
22 // OTHER INFORMATION: n = A,T,C or G
23 //S-10-506-549-3

```

Query Match	1.5%;	Score 36.6;	DB 6;	Length 394191;
Best Local Similarity	44.1%;	Pred. No. 33;		
Matches 153;	Conservative	0;	Mismatches 194;	Indels 0.

1852 TAGAGGGGAATACAAAGGAAGAAAACATTCATTTTATTTGATTCGATTAATCTTAAG 1911
 72108 TACGGGTGTGAACCAACAGCGCTAGCCGCTTTGTATATTTTAATGTATGTTAACTTAAAG 7216

1912 CATTAGCTAATTCAGTAGAGGCATTCACGAGAATTTAAATAGAAATTATATGTACGG 1971

72168 AATAATGATGTTTATATATTCGCTTTTTCACTTAATAATATGTAGAGAACTTTTAAAT

1972 AATATTATTTTGATAGACTGTTGAAAAATTACACAGAGGAAATTGCTGCTCCAG 2031

```
Db 72228 TAGAATTTGTGGCAAACTTTTAAAGACACACCAAGCTTTTATAGATAACT 72287
Qy 2032 TTTTGACAGACACATGATTTGAGTCATTTTAACAGCTAGTCCTTTAACTGT 2091
Db 72288 GTAAATGGTGAATCTCTTTAGCTCAAAATATTTAGCTTTTGCTATTTTCCATACC 72347
Qy 2092 ACAACTGCTGTATATGATGATGATTAACATTAATCTAGTGTGATAGTATT 2151
Db 72348 ATTGTTAATAAGATATTGTGTGTAATCTTTGCACTTTATGATTTTGAATAAAGTA 72407
Qy 2152 ACAGGCTGACACCTGCTCAGTGAAGGTGAGAAAGAGTAAGACTCT 2198
Db 72408 TAAGACTAAATGATTAATTTTAAAGCTGTGACAAATATCAAGCT 72454

RESULT 6
US-11-217-529-4068/c
; Sequence 4068, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4068
; LENGTH: 4209
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-4068

Query Match 1.5%; Score 35.8; DB 7; Length 4209;
Best Local Similarity 49.2%; Pred. No. 9.4;
Matches 94; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

Qy 1877 CATTCGATTTTATTTGATTCGATATCTTATGCAATTTAGCTAATCCAGTAGAGCA 1936
Db 1867 CATTCCTCTTATTTTACTTGTGTGATATTTGCTGAGTTCCCTCTCACTGTAAACCCC 1808
Qy 1937 TTCAGCAGAAATTTAATAGATTAATATGTAAGATATTATTTGATTAAGACTGTTG 1996
Db 1807 TTTCAGTTTCTTCACGTTATTAATTAATCTTACATGATGACATGAGGTGAGGCCAA 1748
Qy 1997 AAAAATTAACAGAGGAAATGCTGCTCCAGTTTGGACAGACACATGATTTGAG 2056
Db 1747 CATACCTACATACGAGGAGTATTCATCTCCAAATTTCCACAGCCGTTTGAATTTAT 1688
Qy 2057 TCATTTTAACA 2067
Db 1687 GTATTTCTCTCA 1677

RESULT 7
US-11-217-529-2925
; Sequence 2925, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
```

```
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2925
; LENGTH: 2286
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-2925

Query Match 1.5%; Score 35.2; DB 7; Length 2286;
Best Local Similarity 53.7%; Pred. No. 10;
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 1839 TGAAGAAGCTCTATAGAGGGAATACAGAGAAAGAAACATTCGATTTTATTTGCATG 1898
Db 1349 TGAGATGCTCGATATAGTATGATTAACAGAAATACCAAGTATATTGTTACTGTTGAGAG 1408
Qy 1899 CGATTAATTTTATGCTTATGCTAATTCGATGAGGCAATTCACAGAAATTTAATAGA 1958
Db 1409 ATTAACACCTTACACATCTGGAATTTGCCGAGAGACTTCCACTACAGAACTCAAGAG 1468
Qy 1959 ATTATATGTAAGAAAT 1974
Db 1469 ATTAATCCGTGTGTAAT 1484

RESULT 8
US-10-473-173-40/c
; Sequence 40, Application US/10473173
; Publication No. US20060088823A1
; GENERAL INFORMATION:
; APPLICANT: VAN ANDEL INSTITUTE
; TITLE OF INVENTION: Microarray Gene Expression Profiling in Clear Cell Renal Cell
; FILE REFERENCE: 38345-170094
; CURRENT APPLICATION NUMBER: US/10/473,173
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US 60/279,411
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 498
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 3323
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-473-173-40

Query Match 1.5%; Score 35.2; DB 6; Length 3323;
Best Local Similarity 57.1%; Pred. No. 12;
Matches 64; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 2217 AAGCTCCTCACTACAGAAAGTCAATTAATAAAGCAACATGATCTTATTTGT 2276
Db 682 AATTTTCATCCACCAAGAAACATCACCAAAATTTCTGGTCAATATGTTTCTTTT 623
Qy 2277 TTTTCTGCTTGATGATTAATGATTAATTTATTTATTTTATTTTATTTTATTTTATTTT 2328
Db 622 CTTTATCTAGAGATCTTTTATACAGATTAATTAAGGTCCGATGATGCTCTGA 571

RESULT 9
US-11-217-529-2163
; Sequence 2163, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
```

```

; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIKO
; FILE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2163
; LENGTH: 1275
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-2163

```

```

Query Match
Best Local Similarity 1.5%; Score 35; DB 7; Length 1275;
Matches 83; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```

```

QY 922 AGGAATGACAGAAACAGACATTCCTGAATGGCTATATATAGAGATACGTAGAG 981
DB 794 AACGACAGAGAAAGAAAGACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 853
QY 982 GTGTCTCTGAATTTAGACTATTTAAAGAGTGAAGACAGATGAGAAATTCATCGCA 1041
DB 854 AAGCCATCGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 913
QY 1042 ATTTCGTAGCTCAGACACTAGACTCGAAGCTTTCTGAAGATCGA 1084
DB 914 AAAAGTCTAGCTCTTCTGAAACGAGATGATGATGAAAGTGA 956

```

```

RESULT 10
US-10-953-349-37222/c
; Sequence 37222, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 37222
; LENGTH: 2439
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-37222

```

```

Query Match
Best Local Similarity 1.5%; Score 35; DB 6; Length 2439;
Matches 62; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
QY 2264 TGTCTATTTTGTCTTCTGCTGATGATTAATGATTAATTAATTAATTAATTAATG 2323
DB 2375 TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 2316
QY 2324 CGTGAATGTGATGAAGAAAGTTAGAGCCACGATATCTGAAG 2370
DB 2315 TATGAATGAGAAATTTGAATGTTGGCTGCTGTATGTACAC 2269

```

```

RESULT 11
US-10-488-619-1326
; Sequence 1326, Application US/10488619
; Publication No. US20060099578A1
; GENERAL INFORMATION:
; APPLICANT: Greenlee, Winner and Sullivan, P.C.
; TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations

```

```

; TITLE OF INVENTION: Physiological Conditions, And Genotyping Arrays
; FILE REFERENCE: 98-01 WO
; CURRENT APPLICATION NUMBER: US/10/488,619
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 3040
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1326
; LENGTH: 111
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-488-619-1326

```

```

Query Match
Best Local Similarity 1.5%; Score 34.6; DB 6; Length 111;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```

```

QY 2261 TGATGTTCTATTTTGTCTTCTGCTGATGATTAATGATTAATTAATTAATTAATTAAT 2320
DB 5 TGTGGGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 64
QY 2321 AGCGTGAATGTGATGAAGAAAGTTAGAGCCAC 2357
DB 65 TTTTCTGGGGGTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 101

```

```

RESULT 12
US-11-315-766-19/c
; Sequence 19, Application US/11315766
; Publication No. US20060101544A1
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene H.
; APPLICANT: Cahoon, Edgar B.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Ramodu, Omolayo O.
; APPLICANT: Hall, Sarah E.
; TITLE OF INVENTION: Phospholipid:diacylglycerol Acetyltransferases
; FILE REFERENCE: B8146 US NA
; CURRENT APPLICATION NUMBER: US/11/315,766
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/321,802
; PRIOR FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 2565
; TYPE: DNA
; ORGANISM: Zea mays
US-11-315-766-19

```

```

Query Match
Best Local Similarity 1.5%; Score 34.6; DB 7; Length 2565;
Matches 52; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 2264 TGTCTATTTTGTCTTCTGCTGATGATTAATGATTAATTAATTAATTAATTAATG 2323
DB 2556 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 2497
QY 2324 CGTGAATGTGATGAAGAAAGAA 2344
DB 2496 TTTTCTGCTCTCAAGAGAA 2476

```

```

RESULT 13
US-10-511-937-646/c
; Sequence 646, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Lv, Ngoc
; APPLICANT: Prentice, James

```

APPLICANT: Morris, Macdonald
APPLICANT: Rosenberg, Steven
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
FILE REFERENCE: 506612000104
CURRENT APPLICATION NUMBER: US/10/511.937
CURRENT FILING DATE: 2004-10-19
PRIOR APPLICATION NUMBER: PCT/US2003/012946
PRIOR FILING DATE: 2003-04-24
PRIOR APPLICATION NUMBER: US 10/131.831
PRIOR FILING DATE: 2002-04-24
PRIOR APPLICATION NUMBER: US 10/325.899
PRIOR FILING DATE: 2002-12-20
NUMBER OF SEQ ID NOS: 3117
SOFTWARE: PatentIn version 3.2
SEQ ID NO 646
LENGTH: 4692
TYPE: DNA
ORGANISM: Homo sapiens
US-10-511-937-646

Query Match 1.5%; Score 34.6; DB 6; Length 4692;
Best Local Similarity 61.8%; Pred. No. 19;
Matches 55; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 2264 TGTCTATTTGTTTCTGCTGATGTTAATGATTATTAATTTTAAAG 2323
DB 4666 TTTTCTATTTTCTGCTGCTGATGTTAATGATTATTAATTTTAAAG 4607
QY 2324 CGTGAATGTGATGAAGAAGATTAGAG 2352
DB 4606 TTTGAACAGAAAGAAAGCTTTTATTAG 4578

RESULT 14
US-10-953-349-6180
Sequence 6180, Application US/10953349
Publication No. US20060107345A1
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nikolai et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
FILE REFERENCE: 2750-1579PUS2
CURRENT APPLICATION NUMBER: US/10/953.349
CURRENT FILING DATE: 2004-09-30
NUMBER OF SEQ ID NOS: 40252
SOFTWARE: PatentIn version 3.3
SEQ ID NO 6180
LENGTH: 624
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-10-953-349-6180

Query Match 1.4%; Score 34.4; DB 6; Length 624;
Best Local Similarity 48.5%; Pred. No. 9.7;
Matches 95; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 1816 TGCACATGCGCTGACATTATTTGAAAAGACTTATAGAGGGGAATCAAGAAAGAAA 1875
DB 420 TGCACATGCTCCCTCTTCTTGAAGTATGATTAATTAATTTCCAAAAAA 479
QY 1876 ACATTGATTTTATTTGATTTGATTAATCTTATGCTTTAGCTTAATTCAGTAGAGC 1935
DB 480 AGTTGGGCTACACACATCAACCTTACTTTTAACTAGACGCAATTCATCTCAA 539
QY 1936 ATTCAGCAGAAATTTAATAGAAATATATAGAAATATTTTGAATAGACTGTTT 1995
DB 540 AAACAAGTAAATCTAACAAGACTAAACATATCTTATCTGTTGGCTGCC 599
QY 1996 GAAAAATTACACAGA 2011
DB 600 TGAATCGAAAGTAGA 615

RESULT 15
US-10-953-349-15742/c
Sequence 15742, Application US/10953349
Publication No. US20060107345A1
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nikolai et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
FILE REFERENCE: 2750-1579PUS2
CURRENT APPLICATION NUMBER: US/10/953.349
CURRENT FILING DATE: 2004-09-30
NUMBER OF SEQ ID NOS: 40252
SOFTWARE: PatentIn version 3.3
SEQ ID NO 15742
LENGTH: 989
TYPE: DNA
ORGANISM: Glycine max
US-10-953-349-15742

Query Match 1.4%; Score 34; DB 6; Length 989;
Best Local Similarity 61.1%; Pred. No. 15;
Matches 55; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 2260 ATGATGTTCTATTTGTTTCTGCTGATGTTAATGATTATTAATTTTAAAG 2319
DB 843 ATTTTCTATTTTCTGCTGCTGATGTTAATGATTATTAATTTTAAAG 784
QY 2320 TAGCGTGAATGTGATGAAGAAGATTAG 2349
DB 783 TTTTCTATTTTCTGCTGCTGATGTTAATGATTATTAATTTTAAAG 754

Search completed: May 27, 2006, 20:04:02
Job time : 43.7971 secs

THIS PAGE BLANK (USPTO)

repeat_region	5402. .5578	/rpt_family="MIR3"
repeat_region	5580. .5637	/rpt_family="(TA)n"
repeat_region		complement(5638. .5713)
repeat_region		/rpt_family="1LMCa"
	6194. .6321	/rpt_family="12"
repeat_region	7681. .7843	/rpt_family="MIR"
repeat_region	7849. .7872	/rpt_family="AT_rich"
repeat_region	8336. .8447	/rpt_family="CT-rich"
repeat_region		complement(8456. .8777)
repeat_region		/rpt_family="A1uSx"
repeat_region		complement(9376. .9682)
repeat_region		/rpt_family="A1uv"
	10920. .11043	
repeat_region		/rpt_family="MER5A"
repeat_region	11353. .11363	/rpt_family="MER5A"
repeat_region		complement(12202. .12268)
repeat_region		/rpt_family="MER5B"
repeat_region	12747. .12861	
repeat_region		/rpt_family="MER81"
repeat_region	13106. .13396	/rpt_family="MER5B"
repeat_region	13397. .13422	
repeat_region		/rpt_family="(GGGGA)n"
repeat_region	13423. .13488	
repeat_region		/rpt_family="MER5B"
repeat_region	14336. .14471	
repeat_region		/rpt_family="MIR"
repeat_region	15241. .15281	/rpt_family="(TG)n"
repeat_region	16465. .16689	
repeat_region		/rpt_family="1LMCa"
repeat_region	16700. .16766	
repeat_region		/rpt_family="(TA)n"
repeat_region	17990. .18044	
repeat_region		/rpt_family="12"
repeat_region	18475. .18600	
repeat_region		/rpt_family="12"
repeat_region	18785. .18937	
repeat_region		/rpt_family="12"
repeat_region		complement(19502. .19270)
repeat_region		/rpt_family="12"
repeat_region	19524. .19676	
repeat_region		/rpt_family="MLT1J2"
repeat_region		complement(19677. .19841)
repeat_region		/rpt_family="12"
repeat_region		complement(120129. .20206)
repeat_region		/rpt_family="MIR"
repeat_region	20779. .21023	
repeat_region		/rpt_family="MIR"
repeat_region		complement(122155. .22575)
repeat_region		/rpt_family="12"
repeat_region	22873. .22987	
repeat_region		/rpt_family="12"
repeat_region	23520. .23568	
repeat_region		/rpt_family="MIR"
repeat_region	25225. .25379	
repeat_region		/rpt_family="MIR3"
repeat_region		complement(125438. .25634)
repeat_region		/rpt_family="MIR"
repeat_region	26260. .26378	
repeat_region		/rpt_family="MIR3"
repeat_region	26592. .26622	
repeat_region		/rpt_family="GC-rich"
repeat_region	28028. .28045	
repeat_region		/rpt_family="(CA)n"
repeat_region	28046. .28077	

Query Match	Beat Local Similarity	9.0%; Score 46; DB 5; Length 163533;
Matches 163; Conservative	0; Mismatches 195; Indels	0; Gaps 0;
repeat_region	28476..28502	/rpf_family="(CA)n"
unsure		/rpf_family="(TTTTA)n"
unsure		complement(28639..28644)
unsure		/note="<30 qual SNGI region"
unsure		28734..28744
unsure		/note="<30 qual SNGI region"
unsure		28768..28774
repeat_region		/note="<30 qual SNGI region"
unsure		28773..28793
unsure		/rpf_family="GC_rich"
unsure		28836..28841
unsure		/note="<30 qual SNGI region"
unsure		28855..28867
repeat_region		/note="<30 qual SNGI region"
repeat_region		28974..29317

Query	23	TTATTATGTGAGTAGAATAGATTCTCCACCAATTAAGATGCAAGACAGAGATTGTG	82
Db	48902	TTAAAGATTGATTAATTTTATCTTAATAATTAAATTATTAACATTTAAATTTANG	48961
Qy	83	TTTATATCTGTGGTAATACGTTTCTCCAGTTGTATTAAGACCTCCACAGATTA	142
Db	48962	GATTACAGATAGATTATATGCTGTATCCAGTGTAAATTTGGTTAACATTTGAA	49021
Qy	143	AGTCCTATGCAACAAAGAAATGCAATACATCTCTAGTCTCATTTATTTTCATTA	202
Db	49022	AATTCATTACTTTTAAATTAATTTTATTTTATTTATTTATTTATTTATTTAGTTGAAC	49081
Qy	203	GATAGCCGGTTTTTACTACACTCAATTAAGATGAACAGATGAATGGTTAGTACTG	262
Db	49082	TTTTTATTTTATTTTATGCGTACATAGAGCTGATATTTAAATGGGATCATGATA	49141
Qy	263	TTTATTAAGAAGATAATAAGATCTATCATCATTTGAGCATTAAGGAGGAGAGAT	322
Db	49142	TATTTGATTCAGACATTAATTAATTTATTTTAAAGATGTTTCAAGCTATTA	49201
Qy	323	TCAGCAACAGTGTGCTTACAGATGGAACAGTTAAAGTAAAGTGAAGCCCCCTCT	380
Db	49202	TTTGTATTTAGTATTAATAATTAAGACATGTTTAAATTAAGTATCTCACCTCTT	49259

Fragment Name	Begin	End	LOCUS	BA000016	Accession	BA000016
Sequence split into 31 fragments						
BA000016_00	1	110000				
BA000016_01	100001	210000				
BA000016_02	200001	310000				
BA000016_03	300001	410000				
BA000016_04	400001	510000				
BA000016_05	500001	610000				
BA000016_06	600001	710000				
BA000016_07	700001	810000				
BA000016_08	800001	910000				
BA000016_09	900001	1010000				
BA000016_10	1000001	1110000				
BA000016_11	1100001	1210000				
BA000016_12	1200001	1310000				
BA000016_13	1300001	1410000				
BA000016_14	1400001	1510000				
BA000016_15	1500001	1610000				
BA000016_16	1600001	1710000				
BA000016_17	1700001	1810000				
BA000016_18	1800001	1910000				
BA000016_19	1900001	2010000				
BA000016_20	2000001	2110000				
BA000016_21	2100001	2210000				

BA000016.22 2200001 2310000
BA000016.23 2300001 2410000
BA000016.24 2400001 2510000
BA000016.25 2500001 2610000
BA000016.26 2600001 2710000
BA000016.27 2700001 2810000
BA000016.28 2800001 2910000
BA000016.29 2900001 3010000
BA000016.30 3000001 3031430
Continuation (15 of 31) of BA000016 from base 1400001 (BA000016 Clostridium perfringens)

Query Match 8.9%; Score 45.6; DB 15; Length 110000;
Best Local Similarity 9.9%; Pred. No. 0.0059;
Matches 105; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 98 AATAAGTTTCTCCAGTGTATTAAGACCCCTCCACAGATTAAGTCCCTATGCAACA 157
DB 15707 AAGTTTTTTTATTAATTTCTTAATATATTAATTAATTAATTAATTAATTAAT 15766
QY 158 AGAAATGCAATACATCTCTAGTCTCATTTATTTTTCATTAGATAGCCGTTTTT 217
DB 15767 AATAGATATTAAGATTAATCTTGACATTAATTAATTAATTAATTAATTAAT 15826
QY 218 ACTACACGTAATTAAGTAACAGATGATGCTTACTGCTTTTAAAGAGAGT 277
DB 15827 TCTTAATATATGATAGGATGCTGGAATGACAAATTAATTAATTAATTAATTA 15886
QY 278 AATTAAGATATCATCATCATTTGA 301
DB 15887 AATTAATACTAATCTCTCAAGA 15910

RESULT 4
ARS79680 1141 bp DNA linear PAT 14-DEC-2004
LOCUS Sequence 22 from patent US 6784342.
ACCESSION ARS79680
VERSION ARS79680.1 GI:56583130
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1141)
AUTHORS Kunze, L. and Clemens, S.
TITLE Regulation of embryonic transcription in plants
JOURNAL Patent: US 6784342-A 22 31-AUG-2004;
The University of British Columbia; Vancouver;
CAN;

FEATURES
Source Location/Qualifiers
1..1141
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 8.9%; Score 45.4; DB 2; Length 1141;
Best Local Similarity 9.9%; Pred. No. 0.01;
Matches 46; Conservative 197; Mismatches 221; Indels 0; Gaps 0;

QY 12 TCATCTCATGCTTATTAATGAGTAAATGATTCACCAATTAAGATGACAAGC 71
DB 110 YBMTNTNKGKTGWRHRYWRABDVHDHYVTAANNATTTGMDKDKRTTMMWKK 169
QY 72 AGAGATTTGTTTATCTGTGGTAATATAGCTTTTCCAGTGTATTAAGACCTGC 131
DB 170 NNATGMDDTKYHMMNNNGCBVTVMVRYKTRDWSBKRMNYGMBWKKMSIDVYYMV 229
QY 132 CACCAAGATTAAGTCCATGCAACAAGAAATGTCAATACATTCCTTAGTCTCATTA 191
DB 230 WDMCKRKYRWVTRGRMRYVVAWBTARRRYNNGTBAAAYRRTVNNNNNNNAKAWC 289
QY 192 TATTTTCATTAGATGAGCGGTTTTTACTACAACTCAATTAAGATTAAGAAATGAG 251
DB 290 KRAKYGMNRABVNSTCTTWSKTTKVRTSCWANNCRAGDANDHKKMKWSAAMGVYNN 349

QY 252 GTTAGAGATGTTTAAAGAGATAATAAGTACTATCATTCATTGAGGCAATAGG 311
DB 350 NNNNNNTYKARHBAKMDVWHSAMKMHANAAYSRKWTBYRKTMVNNNGTTMMKR 409
QY 312 GAGGAGAGATTCAGCAACAGTGTCTTACAGTGAAGAAACAAGTAAGTAAGTAC 371
DB 410 MMATYMKMDMDWBGTYNNNNNGRTYYGWTNKKMMYYTYKKAANCKYRAMDKTCTHN 469
QY 372 CCCCCTCTTACCAAGATCAATGACCAAGTTGAGCTTTAGCCAGCCACATCATGTAA 431
DB 470 TTTMKKTMMNCYKSKMTNGKSHRBAAYTYTMMWRHAYAAANNNDYMKACITWYK 529
QY 432 ATTGCTTCCGATAGACCTGTTCATTAATTCCTTTGCAAGC 475
DB 530 YBVCCKMMNNYAAAYTKSSWNTSRYYRWKRTNNNSMKRSJTRSM 573

RESULT 5
AX083744 1141 bp DNA linear PAT 28-FEB-2001
LOCUS AX083744
DEFINITION Sequence 22 from patent WO0111061.
ACCESSION AX083744
VERSION AX083744.1 GI:13185472
KEYWORDS
SOURCE Synthetic construct
ORGANISM Other sequences; artificial sequences.
REFERENCE 1
AUTHORS Kunze, L. and Clemens, S.
TITLE Regulation of embryonic transcription in plants
JOURNAL Patent: WO 011061-A 22 15-FEB-2001;
UNIVERSITY OF BRITISH COLUMBIA (CA)

FEATURES
Source Location/Qualifiers
1..1141
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
1..1141
/note="consensus sequence of A.C., L.A., and B.N. FAEI
promoters"

ORIGIN
Query Match 8.9%; Score 45.4; DB 2; Length 1141;
Best Local Similarity 9.9%; Pred. No. 0.01;
Matches 46; Conservative 197; Mismatches 221; Indels 0; Gaps 0;

QY 12 TCATCTCATGCTTATTAATGAGTAAATGATTCACCAATTAAGATGACAAGC 71
DB 110 YBMTNTNKGKTGWRHRYWRABDVHDHYVTAANNATTTGMDKDKRTTMMWKK 169
QY 72 AGAGATTTGTTTATCTGTGGTAATATAGCTTTTCCAGTGTATTAAGACCTGC 131
DB 170 NNATGMDDTKYHMMNNNGCBVTVMVRYKTRDWSBKRMNYGMBWKKMSIDVYYMV 229
QY 132 CACCAAGATTAAGTCCATGCAACAAGAAATGTCAATACATTCCTTAGTCTCATTA 191
DB 230 WDMCKRKYRWVTRGRMRYVVAWBTARRRYNNGTBAAAYRRTVNNNNNNNAKAWC 289
QY 192 TATTTTCATTAGATGAGCGGTTTTTACTACAACTCAATTAAGATTAAGAAATGAG 251
DB 290 KRAKYGMNRABVNSTCTTWSKTTKVRTSCWANNCRAGDANDHKKMKWSAAMGVYNN 349
QY 252 GTTAGAGATGTTTAAAGAGATAATAAGTACTATCATTCATTGAGGCAATAGG 311
DB 350 NNNNNNTYKARHBAKMDVWHSAMKMHANAAYSRKWTBYRKTMVNNNGTTMMKR 409
QY 312 GAGGAGAGATTCAGCAACAGTGTCTTACAGTGAAGAAACAAGTAAGTAAGTAC 371
DB 410 MMATYMKMDMDWBGTYNNNNNGRTYYGWTNKKMMYYTYKKAANCKYRAMDKTCTHN 469
QY 372 CCCCCTCTTACCAAGATCAATGACCAAGTTGAGCTTTAGCCAGCCACATCATGTAA 431

Db 470 TTWMMKTYNNCKWKSNTGSKSHRBAALVYTYMMWRRYAHANNNDWYMKACTWYK 529
Oy 432 ATTCCTTTCCTGATAGCCTGTCATTAATCTCTTGCAGAC 475
Db 530 YBVCCKMNNYAAWTKSSWNTSTRYWTKTNSRMSDTRSM 573

RESULT 6
AC094376/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-3Pl. *** SEQUENCING IN PROGRESS ***
AC094376 238301 bp DNA linear HTG 13-NOV-2002
AC094376 2 unordered pieces.
AC094376
AC094376.11 GI:24942492
HTG, HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Rattus.
1 (bases 1 to 238301)
Muzny, D., Maric, M., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Ayalabeche, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Blawie, K., Blair, J., Blankenburg, K., Biyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgescu, E., Geor, K., Gill, R., Garcia, A., Garner, T., Garza, M.,
Gunsberg, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlik, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howell, S., Hulyk, S., Hume, J., Idledid, D., Jackson, A.,
Jackson, B., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Kapathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, M., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenz, L., Louised, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokemele, O., Okunolu, G., Olampunsgoon, A., Pal, S., Parks, K.,
Pastermak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Polndexter, A., Popovic, D., Primus, E., Pu, L., L.,
Piaz, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smales, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorrell, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Swatek, A., Tabors, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, K.,
Williams, G., Wilson, R., Wleczky, R., Woodson, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhou, S., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinrock, G., and Gibbs, R. A.
Direct Submission
2 (bases 1 to 238301)

TITLE
JOURNAL
REFERENCE

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Worley, K. C.
Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 238301)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 13, 2002 this sequence version replaced gi:23265811.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/atlantis/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GAGE
Center clone name: CH230-3Pl
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 229728 bases at least Q40
Consensus quality: 232177 bases at least Q30
Consensus quality: 233540 bases at least Q20
Estimated insert size: 237791; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 237092: contig of 237092 bp in length
* 237093 237192: gap of unknown length
* 237193 238301: contig of 1109 bp in length.
Location/Qualifiers
1. 238301
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-3Pl"
237093..237192
/estimated_length=unknown

ORIGIN
Query Match 8.9%; Score 45.4; DB 12; Length 238301;
Best Local Similarity 53.0%; Pred. No. 0.0064;
Matches 97; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Oy 173 ATTCCTTTCCTGATAGCCTGTCATTAATCTCTTGCAGAC 232
Db 152787 ATTCCTTTCCTGATAGCCTGTCATTAATCTCTTGCAGAC 152728
Oy 233 AGATGAACAGATGAATGCTAGTCTTTTAAAGAGATTAAGATATATC 292

TITLE Homo sapiens, clone RP11-24F22
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 69335)
AUTHORS Birren, B., Linton, L., Nuebaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Balaban, J., Barna, N., Baetien, V., Bede, P.,
Boguslavsky, L., Bouknight, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S.,
Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-pierre, N., Grant, C., Hagos, B., Heatford, A., Horton, L.,
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Laroque, K., Lamazares, R., Landers, T., Lehocsky, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,
McCarthy, M., McGowan, P., McGurk, A., McKernan, K., McPherson, R.,
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenka, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testa, S., Theodore, J., Tjirell, A., Traversa, M., Trigglio, J.,
Vasilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L4446
Center clone name: 24_F_22

* NOTE: This record contains 88 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 709: contig of 709 bp in length
* 710 809: gap of 100 bp
* 810 1511: contig of 702 bp in length
* 1512 1611: gap of 100 bp
* 1612 2309: contig of 698 bp in length
* 2310 2409: gap of 100 bp
* 2410 3108: contig of 699 bp in length
* 3109 3208: gap of 100 bp
* 3209 3893: contig of 685 bp in length
* 3894 3993: gap of 100 bp
* 3994 4686: contig of 693 bp in length
* 4687 4786: gap of 100 bp
* 4787 5482: contig of 696 bp in length
* 5483 5582: gap of 100 bp
* 5583 6292: contig of 710 bp in length
* 6293 6392: gap of 100 bp
* 6393 7092: contig of 700 bp in length
* 7093 7192: gap of 100 bp
* 7193 7865: contig of 673 bp in length
* 7866 7965: gap of 100 bp
* 7966 8657: contig of 692 bp in length
* 8658 8757: gap of 100 bp
* 8758 9444: contig of 687 bp in length

9445 9544: gap of 100 bp
* 9545 10237: contig of 693 bp in length
* 10238 10337: gap of 100 bp
* 10338 11014: contig of 677 bp in length
* 11035 11114: gap of 100 bp
* 11115 11778: contig of 664 bp in length
* 11779 11878: gap of 100 bp
* 11879 12555: contig of 677 bp in length
* 12556 12655: gap of 100 bp
* 12656 13360: contig of 705 bp in length
* 13361 13460: gap of 100 bp
* 13461 14163: contig of 703 bp in length
* 14164 14263: gap of 100 bp
* 14264 14934: contig of 671 bp in length
* 14935 15034: gap of 100 bp
* 15035 15720: contig of 686 bp in length
* 15721 15820: gap of 100 bp
* 15821 16514: contig of 694 bp in length
* 16515 16614: gap of 100 bp
* 16615 17304: contig of 690 bp in length
* 17305 17404: gap of 100 bp
* 17405 18093: contig of 689 bp in length
* 18094 18193: gap of 100 bp
* 18194 18891: contig of 698 bp in length
* 18892 18991: gap of 100 bp
* 18992 19672: contig of 681 bp in length
* 19673 19772: gap of 100 bp
* 19773 20455: contig of 683 bp in length
* 20456 20555: gap of 100 bp
* 20556 21339: contig of 684 bp in length
* 21340 22026: contig of 687 bp in length
* 22027 22126: gap of 100 bp
* 22127 22826: contig of 700 bp in length
* 22827 22926: gap of 100 bp
* 22927 23625: contig of 699 bp in length
* 23626 23725: gap of 100 bp
* 23726 24393: contig of 668 bp in length
* 24394 24493: gap of 100 bp
* 24494 25181: contig of 688 bp in length
* 25182 25281: gap of 100 bp
* 25282 25984: contig of 703 bp in length
* 25985 26084: gap of 100 bp
* 26085 26765: contig of 681 bp in length
* 26766 27573: gap of 100 bp
* 27574 27673: contig of 708 bp in length
* 27674 28453: contig of 680 bp in length
* 28454 29142: contig of 689 bp in length
* 29143 29242: gap of 100 bp
* 29243 29927: contig of 685 bp in length
* 29928 30027: gap of 100 bp
* 30028 30719: contig of 692 bp in length
* 30720 31514: contig of 695 bp in length
* 31515 31614: gap of 100 bp
* 31615 32308: contig of 694 bp in length
* 32309 32408: gap of 100 bp
* 32409 33094: contig of 686 bp in length
* 33095 33194: gap of 100 bp
* 33195 33895: contig of 701 bp in length
* 33896 33995: gap of 100 bp
* 33996 34705: contig of 710 bp in length
* 34706 34805: gap of 100 bp
* 34806 35504: contig of 699 bp in length
* 35505 35604: gap of 100 bp
* 35605 36291: contig of 687 bp in length
* 36292 36391: gap of 100 bp
* 36392 37090: contig of 696 bp in length
* 37091 37190: gap of 100 bp
* 37191 37881: contig of 691 bp in length
* 37882 37981: gap of 100 bp

```

*      37982      38671: contig of 690 bp in length
*      38672      38771: gap of 100 bp
*      38772      39465: contig of 694 bp in length
*      39466      39565: gap of 100 bp
*      39566      40255: contig of 690 bp in length
*      40256      40355: gap of 100 bp
*      40356      41052: contig of 697 bp in length
*      41053      41152: gap of 100 bp
*      41153      41847: contig of 695 bp in length
*      41848      41947: gap of 100 bp
*      41948      42630: contig of 683 bp in length
*      42631      42731: gap of 100 bp
*      42731      43417: contig of 687 bp in length
*      43418      43517: gap of 100 bp
*      43518      44185: contig of 668 bp in length
*      44186      44285: gap of 100 bp
*      44286      44986: contig of 701 bp in length
*      44987      45086: gap of 100 bp
*      45087      45766: contig of 680 bp in length
*      45767      45866: gap of 100 bp
*      45867      46530: contig of 664 bp in length
*      46531      46631: gap of 100 bp
*      46631      47323: contig of 693 bp in length
*      47324      47423: gap of 100 bp
*      47424      48022: contig of 599 bp in length
*      48023      48122: gap of 100 bp
*      48123      48778: contig of 656 bp in length
*      48779      48878: gap of 100 bp
*      48879      49579: contig of 701 bp in length
*      49580      49679: gap of 100 bp
*      49680      50387: contig of 708 bp in length
*      50388      50487: gap of 100 bp
*      50488      51192: contig of 705 bp in length
*      51193      51292: gap of 100 bp
*      51293      51980: contig of 688 bp in length
*      51981      52080: gap of 100 bp
*      52081      52784: contig of 704 bp in length
*      52785      52884: gap of 100 bp
*      52885      53572: contig of 688 bp in length
*      53573      53672: gap of 100 bp
*      53673      54365: contig of 693 bp in length
*      54366      54465: gap of 100 bp

```

Query Match Best Local Similarity 8.7%; Score 44.6; DB 12; Length 69335; Matches 145; Conservative 0; Mismatches 149; Indels 1; Gaps 1;

```

Qy      49  CACCAATTTAGATGACAAAGACAGATTTGTTTATCTGTTGGTAAATACGTTT 108
Db      62189 CAGGCACACGCGGACATGTATATCATATGTAACCTGCAACATGTGCACATGTA 62130
Qy      109  CTCGAGTTGTAAGAACCCCTCCACCACTATTAAGTCTTATGCAACAAAGAAATGTC 168
Db      62129 CCTTAAACCTTAAAGTATTAATTAATTAATTAATTAATTAATTAATTAAT 62070
Qy      169  ATACATTTCTTATGTCATATATTTTTCATAGATAGCCGTTTTCCTCACTCA 228
Db      62069 TAAATTTCTTAAATCT-ACCTTCAGTGAATCTCGAAGCTTTTCTTAAATTAATGA 62011
Qy      229  AATTAAGTAAACAGATGAATGAGTTAGTGACTGTTTATTAAGAAAGATTAAGATAC 288
Db      62010 CTTATATTTTCAACAGTGTATGAATATGCGCTATTTTAAAGAAACAGTATTTATTTG 61951
Qy      289  TATCATATTTGAGGCAATAGAGAGAGAGATTCAGCAACAGTGTCTTCA 343
Db      61950 AAAAAACCTTATGATGTGAAGAGAGAGAGATGAAACCAAGCCTGTATATAA 61896

```

RESULT 9 AC127531 52915 bp DNA linear HTG 17-JUN-2002
 AC127531 Homo sapiens chromosome 11 clone CTD-2362F20 map 11, LOW-PASS
 DEFINITION SEQUENCE SAMPLING.

ACCESSION AC127531
 VERSION AC127531.1 GI:21886951
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 52915)
 Birren, B., Nusbaum, C. and Lander, E.
 Homo sapiens chromosome 11, clone CTD-2362F20
 REFERENCE 2 (bases 1 to 52915)
 Unpublished

REFERENCE 1 (bases 1 to 52915)
 Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Batra, N., Baerlein, V., Bloom, T., Boguslavsky, L., Boulhagalter, B., Camarata, J., Chang, J., Chazaro, B., Chesepel, Y., Collymore, A., Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardina, S., Gord, S., Graham, L., Grand-pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Keller, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Mathews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Plunkhang, P., Pierre, N., Raymond, C., Retz, R., Rise, C., Rogov, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vasiliyev, H., Vei, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (17-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L27588
 Center clone name: 2362_F_20

* NOTE: This record contains 64 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.
 * 1
 * 689 688: contig of 688 bp in length
 * 789 788: gap of 100 bp
 * 1509 1509: contig of 721 bp in length
 * 1510 1509: gap of 100 bp
 * 1610 2354: contig of 745 bp in length
 * 2355 2454: gap of 100 bp
 * 2455 3158: contig of 704 bp in length
 * 3159 3258: gap of 100 bp
 * 3259 3989: contig of 731 bp in length
 * 3990 4090: gap of 100 bp
 * 4090 4825: contig of 736 bp in length
 * 4826 4925: gap of 100 bp
 * 4926 5657: contig of 732 bp in length
 * 5658 5757: gap of 100 bp
 * 5758 6490: contig of 732 bp in length
 * 6490 6589: gap of 100 bp

```
* 6590 7298: contig of 709 bp in length
* 7299 7398: gap of 100 bp
* 7399 8132: contig of 734 bp in length
* 8133 8232: gap of 100 bp
* 8233 8962: contig of 730 bp in length
* 8963 9062: gap of 100 bp
* 9063 9790: contig of 728 bp in length
* 9791 9890: gap of 100 bp
* 9891 10625: contig of 735 bp in length
* 10626 10725: gap of 100 bp
* 10726 11456: contig of 731 bp in length
* 11457 11556: gap of 100 bp
* 11557 12283: contig of 727 bp in length
* 12284 13130: contig of 747 bp in length
* 13131 13230: gap of 100 bp
* 13231 13983: contig of 753 bp in length
* 13984 14083: gap of 100 bp
* 14084 14832: contig of 749 bp in length
* 14833 14932: gap of 100 bp
* 14933 15661: contig of 729 bp in length
* 15662 15761: gap of 100 bp
* 15762 16482: contig of 721 bp in length
* 16483 16582: gap of 100 bp
* 16583 17298: contig of 716 bp in length
* 17299 17398: gap of 100 bp
* 17399 18144: contig of 746 bp in length
* 18145 18244: gap of 100 bp
* 18245 18979: contig of 735 bp in length
* 18980 19079: gap of 100 bp
* 19080 19797: contig of 718 bp in length
* 19798 19897: gap of 100 bp
* 19898 20621: contig of 724 bp in length
* 20622 20721: gap of 100 bp
* 20722 21470: contig of 749 bp in length
* 21471 21570: gap of 100 bp
* 21571 22316: contig of 746 bp in length
* 22317 22416: gap of 100 bp
* 22417 23125: contig of 709 bp in length
* 23126 23225: gap of 100 bp
* 23226 23968: contig of 743 bp in length
* 23969 24068: gap of 100 bp
* 24069 24794: contig of 726 bp in length
* 24795 24894: gap of 100 bp
* 24895 25621: contig of 727 bp in length
* 25622 25721: gap of 100 bp
* 25722 26453: contig of 732 bp in length
* 26454 26553: gap of 100 bp
* 26554 27300: contig of 747 bp in length
* 27301 27400: gap of 100 bp
* 27401 28125: contig of 725 bp in length
* 28126 28225: gap of 100 bp
* 28226 28974: contig of 749 bp in length
* 28975 29074: gap of 100 bp
* 29075 29816: contig of 742 bp in length
* 29817 29916: gap of 100 bp
* 29917 30663: contig of 747 bp in length
* 30664 30763: gap of 100 bp
* 30764 31497: contig of 734 bp in length
* 31498 31597: gap of 100 bp
* 31598 32331: contig of 734 bp in length
* 32332 32431: gap of 100 bp
* 32432 33159: contig of 728 bp in length
* 33160 33259: gap of 100 bp
* 33260 33988: contig of 729 bp in length
* 33989 34088: gap of 100 bp
* 34089 34835: contig of 747 bp in length
* 34836 34935: gap of 100 bp
* 34936 35671: contig of 736 bp in length
* 35672 35771: gap of 100 bp
* 35772 36505: contig of 734 bp in length
* 36506 37286: gap of 100 bp
* 37286: contig of 681 bp in length
```

```
* 37287 37386: gap of 100 bp
* 37387 38115: contig of 729 bp in length
* 38116 38215: gap of 100 bp
* 38216 38944: contig of 729 bp in length
* 38945 39045: gap of 100 bp
* 39045 39765: contig of 721 bp in length
* 39766 39865: gap of 100 bp
* 39866 40599: contig of 734 bp in length
* 40600 40699: gap of 100 bp
* 40700 41441: contig of 742 bp in length
* 41442 41541: gap of 100 bp
* 41542 42157: contig of 616 bp in length
* 42158 42257: gap of 100 bp
* 42258 42989: contig of 732 bp in length
* 42990 43089: gap of 100 bp
* 43090 43805: contig of 716 bp in length
* 43806 43905: gap of 100 bp
* 43906 44654: contig of 749 bp in length
* 44655 44754: gap of 100 bp
* 44755 45498: contig of 744 bp in length
* 45499 45598: gap of 100 bp
* 45599 46335: contig of 737 bp in length
* 46336 46435: gap of 100 bp
* 46436 47149: contig of 714 bp in length
* 47150 47249: gap of 100 bp
* 47250 47980: contig of 731 bp in length
* 47981 48080: gap of 100 bp
* 48081 48819: contig of 739 bp in length
* 48820 48919: gap of 100 bp
* 48920 49646: contig of 727 bp in length
* 49647 49746: gap of 100 bp
* 49747 50473: contig of 727 bp in length
* 50474 50573: gap of 100 bp
* 50574 51269: contig of 695 bp in length
* 51269 51368: gap of 100 bp
* 51369 52094: contig of 726 bp in length
* 52095 52194: gap of 100 bp
* 52195 52915: contig of 721 bp in length.

FEATURES
source
1..52915
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="11"
    /map="11"
    /clone="CTD-2362F20"
    /clone_id="CITD1 Human BAC"
    /estimated_length=100
    /estimated_length=100
    /estimated_length=100

Query Match      8.7%; Score 44.4; DB 12; Length 52915;
Best Local Similarity 50.0%; Pred.No. 0.015;
Matches 141; Conservative 0; Mismatches 136; Indels 5; Gaps 1;

QY 19 ATGCTTATATGAGTGAATAGTATCTTCACCAAAATTGAATGACAAAGCAAGAT 78
DB 39896 ATTCGCGAGTGAGATGACAAATCTTAAGCAAAAGTAAAGTCAGAGAGAT 39955
QY 79 TGTGTTTATCTGTGGTAAATAGTTTCTCCAGTTGTATTAAGACCTCCACGAGT 138
DB 39956 AAATTCAGATCTATCTACTAATATTTCTATCAAGAACTGGGAAAACTTCCA----- 40010
QY 139 ATAAAGTCATGCAACAAAGAAATGCAATTAATCTCTTAACTCATTTATTTTC 198
DB 40011 AATAAGTCACCTGAACATAAATAATTAGAAATATGTTGTAATAAGTGAAGTCTT 40070
QY 199 ATTGATATGCGCGTTTCTTACTACATCAATTAAGATGAACAGATGAATGGTTAGT 258
DB 40071 AAAAGATGAGATGAGAACTACCGCTTCAGTCTGAGGAAGTGAAGTGAACAA 40130
QY 259 ACTGTTTATAAGAAAGTAAATAAAGATATCATTCATTTG 300
```

Db 40131 ACCTCAAACTACAGAGAAAGTGAACATGTTGTAACCTTGG 40172

RESULT 10
CR956401 206489 bp DNA linear HTG 30-JAN-2006
LOCUS Danio rerio clone DKEX-5711, WORKING DRAFT SEQUENCE, 11 unordered
DEFINITION
piece.
ACCESSION CR956401.3 GI:86197649
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULUTOP.
KEYWORDS Danio rerio (zebrafish)
SOURCE
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Osteichthyes;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 206489)
McIaren,S.
Direct Submission
Submitted (29-JAN-2006) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests:
http://www.sanger.ac.uk/Projects/D_rerio/fags.shtml#dataeight
On Jan 30, 2006 this sequence version replaced gi:67509286.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: ZK5711
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 202739 bases at least Q40
Consensus quality: 203666 bases at least Q30
Consensus quality: 204251 bases at least Q20
Insert size: 205489; sum-of-contigs
Insert size: 249713; 4.3% error; agarose-fp
Quality coverage: 12.96x in Q20 bases; sum-of-contigs Quality
coverage: 10.86x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 36677: contig of 36677 bp in length
* 36678 36777: gap of 100 bp
* 36778 79240: contig of 42463 bp in length
* 79241 79340: gap of 100 bp
* 79341 86787: contig of 7447 bp in length
* 86788 86888: gap of 100 bp
* 86888 93958: contig of 7071 bp in length
* 93959 94058: gap of 100 bp
* 94059 101325: contig of 7267 bp in length
* 101326 101425: gap of 100 bp
* 101426 106157: contig of 4732 bp in length
* 106158 106257: gap of 100 bp
* 106258 110350: contig of 4103 bp in length
* 110351 110460: gap of 100 bp
* 110461 116362: contig of 5902 bp in length
* 116363 116462: gap of 100 bp
* 116463 176493: contig of 60031 bp in length
* 176494 176593: gap of 100 bp
* 176594 180881: contig of 4288 bp in length
* 180882 180982: gap of 100 bp
* 180983 206489: contig of 25508 bp in length.
Location/Qualifiers

source 1..206489
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEX-5711"
/clone_1b="DanioKey"
1..36677
/note="assembly fragment:00001
fragment chain:1"
36778..79240
/note="assembly fragment:00713
fragment chain:1"
79341..86787
/note="assembly fragment:03265
fragment chain:1"
86888..93958
/note="assembly fragment:03188
fragment chain:1"
94059..101325
/note="assembly fragment:02996
fragment chain:1"
101426..106157
/note="assembly fragment:02967
fragment chain:1"
106258..110350
/note="assembly fragment:03032
fragment chain:1"
110461..116362
/note="assembly fragment:03077
fragment chain:1"
116463..176493
/note="assembly fragment:01663"
176594..180881
/note="assembly fragment:03130.0"
180982..206489
/note="assembly fragment:03345"

ORIGIN
Query Match 8.7%; Score 44.4; DB 12; Length 206489;
Best Local Similarity 43.5%; Pred. No. 0.013;
Matches 117; Conservative 0; Mismatches 152; Indels 0; Gaps 0;
QY 55 ATTAGATGACAAAGAGAGATTGCTTTATCTGCTGGTAAATACGTTCTCAG 114
DB 110503 AATTAATAAAAAAAAAAANTTTNTNTTTTAAAAAATAATATTTTNNN 110562
QY 115 TTGATTAAGACCCCTCCACAGATTAAGTCTATGCAACAAAGAAATGCAATACAT 174
DB 110563 NNN 110622
QY 175 TCTCTTACTGCTATTAATTAATTTCAATTAAGTACCGGTTTCTACAACTCAATAG 234
DB 110623 ATATTTTATTTTATTTTATTTTAAANNNTTATNTTTTATTTTATTTTAA 110682
QY 235 ATGAACAGAAAGATGCGTTAGTGCCTTTATTAAGAGATTAAGATTAATCAT 294
DB 110683 AAAAAAAAAAATTAATTAATTTTATTTTAAATATTTTAAAAAATTTT 110742
QY 295 CATTGAGCAATTAAGGAGAGAGATTT 323
DB 110743 TTTATATTAATAAATAATATATATATT 110771

RESULT 11
PFMAL13PB/C 83110 bp DNA linear HTG 11-AUG-1999
LOCUS Plasmodium falciparum 3D7 chromosome 13, *** SEQUENCING IN PROGRESS
DEFINITION *** 9 unordered pieces.
ACCESSION AL109814.1 GI:5731891
VERSION HTG; HTGS_PHASE1.
KEYWORDS Plasmodium falciparum 3D7
SOURCE
ORGANISM Plasmodium falciparum 3D7

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	Bowman, S., Churcher, C., Harris, B., Harris, D., Lawson, D., Quail, M. and Barrett, B.	Submitted (11-AUG-1999) P.falciparum Genome Sequencing Consortium, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK	For more information about this sequence or the Malaria Project, visit http://www.sanger.ac.uk/Projects/Pfalciparum/

PFMAL13_14
WPCOMMENT

Sequence split into 28 fragments LOCUS PFMAL13 Accession AL844509

Fragment Name	Begin	End
PFMAL13_00	1	110000
PFMAL13_01	100001	210000
PFMAL13_02	200001	310000
PFMAL13_03	300001	410000
PFMAL13_04	400001	510000
PFMAL13_05	500001	610000
PFMAL13_06	600001	710000
PFMAL13_07	700001	810000
PFMAL13_08	800001	910000
PFMAL13_09	900001	1010000
PFMAL13_10	1000001	1100000
PFMAL13_11	1100001	1210000
PFMAL13_12	1200001	1310000
PFMAL13_13	1300001	1410000
PFMAL13_14	1400001	1510000
PFMAL13_15	1500001	1610000
PFMAL13_16	1600001	1710000
PFMAL13_17	1700001	1810000
PFMAL13_18	1800001	1910000
PFMAL13_19	1900001	2010000
PFMAL13_20	2000001	2100000
PFMAL13_21	2100001	2210000
PFMAL13_22	2200001	2310000
PFMAL13_23	2300001	2410000
PFMAL13_24	2400001	2510000
PFMAL13_25	2500001	2610000
PFMAL13_26	2600001	2710000
PFMAL13_27	2700001	2732359

Continuation (15 of 28) of PFMAL13 from base 1400001 (AL844509 Plasmodium falciparum 3D7)

Query Match 8.6%; Score 43.8; DB 12; Length 110000;
 Best Local Similarity 46.3%; Pred. No. 0.022;
 Matches 144; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

Qy	4	TGATCCTTCATCTCATCTATTTATGCGAGTAAATGATTCACCAATATGAAG	63
Db	3932	TTGTTTTTCTTCAACAAATGTAATTAATTAATTTTACTATATTAATTTTA	3991
Qy	64	GACAAAGACAGATTTGTTTATCTGTGGTAATACGTTTCTCCAGTTGATATA	123
Db	3992	AACAAATTAATTTTATTTTATTAAGATTAATAAATATATGTTTACATATA	4051
Qy	124	GACCTCCACAGATATAAGTCTATGCAACAAAGAAATGTCATACATCTCTT	183
Db	4052	AATCTATATATGATTAATAAATCTATGACACCTTAATTTAGATTAATTTT	4111
Qy	184	CTCATATATTTTCAATTAAGTGGGTTTCTCACTCAATTAAGTGAACAGA	243
Db	4112	TTTTTTTATCTTAATTTTTCGTCATTTTATTAATTAACATAAATAAATAA	4171
Qy	244	ATGAATGGGTTAGTACTGTTTATTAAGAGATTAATTAAGTACTATCATTT	303
Db	4172	AAAAAATATCAAAAGAAATGAATAATTAATTAATTAATAAAGAACCGTT	4231
Qy	304	CAATTAAGGAG	314
Db	4232	ATTAAGGAGCG	4242

RESULT 14
 AL136371 147784 bp DNA linear PRI 18-MAY-2005
 LOCUS Human DNA sequence from clone RP11-110P20 on chromosome 1q31.1
 DEFINITION Contains STSs, GSSs and a Cpg island, complete sequence.
 ACCESSION AL136371
 VERSION AL136371.8 GI:9187156
 KEYWORDS HTG; Cpg island.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 147784)
 Chapman, J.
 Direct Submission
 Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
 Clone requests: clonerequest@sanger.ac.uk
 On Jul 14, 2000 this sequence version replaced gi:8894185.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
 on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
 was generated from part of bacterial clone contigs of human
 chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
 Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr1>
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: vegas@sanger.ac.uk

RP11-110P20 is from the library RPI1-11.1 constructed by the group
 of Pieter de Jong. For further details see
<http://www.choi.org/bacpac/home.htm>
 VECTOR: pBAC3.6
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one subclone; and the assembly was confirmed by restriction digest,
 except on the rare occasion of the clone being a YAC.

FEATURES
source

1..147784
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="q31.1"
 /clone="RP11-110P20"
 /clone_1id="RPI1-11.1"
 1
 /note="Clone_left_end: RP11-110P20"
 59783
 /note="Clone_left_end: RP11-487023"
 147784
 /note="Clone_right_end: RP11-110P20"

ORIGIN

Query Match 8.6%; Score 43.8; DB 5; Length 147784;
 Best Local Similarity 53.8%; Pred. No. 0.021;
 Matches 133; Conservative 0; Mismatches 112; Indels 2; Gaps 2;

Qy	39	AATAGATTCTCACCATAATGATGACAAAGAGATTTGTTTATCTGTTGGGTA	98
Db	35138	ACTTCACCTTTACCAAGTTAATAATTAATTAATTTGTGAAGATTTTGTCTTTTA	35197
Qy	99	AATAGTTTCTCCAGTTGATA-AAGACCTCCACAGATATAAGTCTTATGACACA	157
Db	35198	AAAGTTAGTTTCTTACTATATAGAGAGTATATAAATTTAAAGTTTATGTTATTA	35257
Qy	158	AGAAATGTCATATCATCTCTTAGTCTCATTTATAT-TTTCATTAAGTACCGGTTT	216
Db	35258	AGGCTTTAAAGAAATAAGTCTTGCCATATATTTATGTTTCATTTCAATCTGTTAG	35317
Qy	217	TACTACACTCAATTAAGATGAACAGATGATGGTTAGTACTGTTTATTAAGAGAG	276
Db	35318	TCCTATATCTCAGTGTAGTTAAAGCCTGAAGAGGTATTTGCAATGTTTGTATGGAAA	35377

QY 277 TAATAAA 283
Db 35378 TCTTAGA 35384

RESULT 15
ABI316154/c
LOCUS
DEFINITION Homo sapiens DNA, STS on chromosome 1, DIS19691, sequence tagged site.
ACCESSION ABI316154 413 bp DNA linear STS 27-OCT-2005
VERSION ABI316154.1 GI:62156735
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Tamiya,G., Shinya,M., Imanishi,T., Ikuta,T., Makino,S., Okamoto,K., Furugaki,K., Matsumoto,T., Mano,S., Ando,S., Nozaki,Y., Yukawa,M., Nakashige,R., Yamaguchi,D., Ishibashi,H., Yonekura,M., Nakami,Y., Takayama,S., Endo,T., Saruwatari,T., Yagura,M., Yoshikawa,Y., Fujimoto,K., Oka,A., Chiku,S., Linzen,S.E., Giphart,M.J., Kulski,J.K., Fukazawa,T., Hashimoto,H., Kimura,M., Hoshina,Y., Suzuki,Y., Hotta,T., Mochida,J., Minezaki,T., Komai,K., Shiozawa,S., Taniguchi,A., Yamanaka,H., Kametani,N., Gojobori,T., Bahram,S. and Inoko,H.
TITLE Whole genome association study of rheumatoid arthritis using 27 039 microsatellites
JOURNAL Hum. Mol. Genet. 14 (16), 2305-2321 (2005)
REFERENCE PUBMED 16000323
AUTHORS 2 (bases 1 to 413)
Tamiya,G., Makino,S., Fujimoto,K., Oka,A., Hayashi,H., Denda,A., Linzen,S.E., Ikuta,T., Shinya,M., Endo,T., Tomizawa,M., Tokudo,E., Sato,R., Takaki,A., Nagatsuka,Y., Watanabe,H., Adachi,S., Makino,Y., Nakano,S., Yamamoto,A., Yoshida,K., Okamoto,K., Yamaguchi,D., Ishibashi,H., Yonekura,M., Takayama,S., Nakami,Y., Saruwatari,T., Brand,A., Van Hiltten,J.A., van de Watering,L.M., Giphart,M.J., Bahram,S., Kulski,Y.J. and Inoko,H.
TITLE Direct Submission
JOURNAL Submitted (12-DEC-2003) Hidetoshi Inoko, Tokai University School of Medicine, Department of Genetic Information, Bohseidai, Isehara, Kanagawa, 259-1193, Japan (E-mail:hinoko@is.icc.u-tokai.ac.jp, Tel:81-463-93-1121, Fax:81-463-94-8884)
COMMENT Polymorphisms were confirmed by comparing with the pooled DNA typing data of 88 Dutch population or 100 Australian. When there was no significant difference between the Japanese and these Caucasians, individual typing was performed to confirm polymorphisms.

FEATURES
SOURCE
1. 413
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="1q43"
/note="pooled DNA of 100 Japanese unrelated individuals sequenced site DIS19691"
1. 24
/note="5' primer: AGTGGTATGATATGCTTCATATGTC"
PCR conditions="denaturation 96degC 5 min, 57degC 1 min, 72degC 1 min; 40 cycles 96degC 45 sec, 57degC 45 sec, 72degC 1 min"
178. 227
/note="sequence tagged site DIS19691"
/rpt_type=cadem
/rpt_unit_seq="aat"
complement(390..413)
/note="3' primer: GTGGAACTGTAGTCCATTAACT"

ORIGIN
Query Match 8.5%; Score 43.6; DB 7; Length 413;

Best Local Similarity 55.2%; Pred. No. 0.04;
Matches 85; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 138 TATAAAGCTATGCAACAAAGAAATGCAATACATCTCTTAGCTCATTTATTTT 197
Db 256 TATAGCCAACTAGCCAACTAGGCTACTTTTATTATTATTATTATTATTATTAT 197
QY 198 CATTAGATAGCCGGTTTTCTTACTACACTCAATATAGATGAAACAGATGAATGGTACT 257
Db 196 TATTATTACTATTATTATTGCTCTCAGCTCCAAAGGCTACAGAAAGCAAGTACTAT 137
QY 258 GACTGTTTATTAAGAGAGTAAATATAGATCAT 291
Db 136 GACTTATATAGTATGAAATATATGATTAATCTAATTT 103

Search completed: May 28, 2006, 10:39:27
UOD time : 2989.22 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2006, 18:26:02 ; Search time 343.339 Seconds
(without alignments)
10397.276 Million cell updates/sec

Title: US-10-099-663-1_COPY_1115_1626

Perfect score: 512
Sequence: 1 cagtgagatccctcattccatc.....gtctgcctcacagacagaag 512

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: N_Geneseq_8:*
- 2: Geneseq1980s:*
- 3: Geneseq1990s:*
- 4: Geneseq2000s:*
- 5: Geneseq2001as:*
- 6: Geneseq2001bs:*
- 7: Geneseq2002as:*
- 8: Geneseq2003as:*
- 9: Geneseq2003bs:*
- 10: Geneseq2003cs:*
- 11: Geneseq2003ds:*
- 12: Geneseq2004as:*
- 13: Geneseq2004bs:*
- 14: Geneseq2005s:*
- 15: Geneseq2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	512	100.0	2381	11	ADL90127 Chicken 1
2	326	63.7	336	11	ADL90128 Chicken 1
3	42.8	8.4	2000	11	ACL35363 Rice stre
4	42.6	8.3	2000	8	ADA71938 Rice gene
5	42.4	8.3	5908	4	AA545387 Chemical1
6	42.4	8.3	5908	6	ABK28232 DNA trans
7	41.8	8.2	10048	6	ABL70313 Chemical1
8	41.8	8.2	10048	6	AA561251 Human gen
9	40.6	7.9	7040	4	AA546439 Tumour su
10	40.6	7.9	7040	6	ABK33963 Human DNA
11	40.6	7.9	7040	8	AB210179 Haematopo
12	40.6	7.9	7040	8	AB210033 Haematopo
13	40.6	7.9	7040	8	ADA20348 Prostate
14	40.6	7.9	7040	8	ADA84155 Human ren
15	40.6	7.9	7040	10	ADA84191 Human lym
16	40.6	7.9	7040	10	ADA84115 Human lym
17	40.6	7.9	14095	6	ABL32477 Human imm
18	40.6	7.9	110000	14	AED76257_3 Continuation (4 of

19	40.2	7.9	6065	6	ABK31356 Signal tr
20	40.2	7.9	6065	6	ABL70579 Chemical1
21	40.2	7.9	6065	6	AA561260 Human gen
22	40.2	7.9	8365	2	AA220056 Plasmodiu
23	39.8	7.8	2000	11	ACL35363 Rice stre
24	39.6	7.7	6086	14	ADZ64438 Human can
25	39.6	7.7	110000	14	ABE04876_4 Continuation (5 of
26	39.4	7.7	5265	6	ABK28419 DNA trans
27	39.4	7.7	110000	10	ACF67367_23 Continuation (24 o
28	39.4	7.7	110000	10	ACF65386_5 Continuation (6 of
29	39.2	7.7	9117	6	ABL33069 Human imm
30	39	7.6	2933	4	ABL18616 Drosoph11
31	39	7.6	5311	6	ABL33019 Human imm
32	39	7.6	6202	13	ADV35113 Human CDN
33	39	7.6	6240	10	ADD49062 Human NOV
34	39	7.6	22052	4	ABL20542 Drosoph11
35	39	7.6	26097	4	ABL03596 Drosoph11
36	39	7.6	36194	4	ABL20582 Drosoph11
37	39	7.6	59967	4	ABL15492 Drosoph11
38	38.4	7.5	32392	6	ABL56203 AMEPV gen
39	38.4	7.5	50000	6	ABL55643 AMEPV gen
40	38.2	7.5	6072	6	ABL32031 Human hyp
41	38.2	7.5	56737	6	AB569895 Human hyp
42	38.2	7.5	56737	15	AEF10056 Human hyp
43	38.2	7.5	110000	12	ADH69807_4 Continuation (5 of
44	38	7.4	2000	11	ACL37108 Rice stre
45	38	7.4	6802	6	ABL33231 Human imm

ALIGNMENTS

RESULT 1

ID ADL90127 standard; DNA; 2381 BP.

ADL90127;

20-MAY-2004 (first entry)

Chicken intestinal fatty acid binding protein, tFABP, gene, 5' region.

Chicken; ds; intestinal fatty acid binding protein; tFABP; gut specific promoter; transgenic.

Gallus gallus.

US2003177516-A1.

18-SEP-2003.

14-MAR-2002; 2002US-00099663.

14-MAR-2002; 2002US-00099663.

(HORS/) HORSEMAN N D.

(PRATT/) PRATT S L.

Horseman ND, Pratt SL;

WPI; 2003-898653/82.

New nucleic acid molecule comprising an isolated avian gut-specific gene expression control region, useful for regulating heterologous nucleic acids in transgenic avians, and for generating transgenic birds.

Claim 1; SEQ ID NO 1; 28pp; English.

The invention relates to an isolated nucleic acid comprising an isolated avian gut-specific gene expression control region appearing as ADL90127 (Chicken intestinal fatty acid binding protein, tFABP, gene, 5' region or ADL90128 (Chicken tFABP promoter) or its degenerate variant. Also included are a recombinant DNA molecule comprising an isolate avian

CC gut-specific gene expression control region operably linked to a nucleic
CC acid insert encoding a polypeptide, an expression vector that integrates
CC into a host cell (and comprising the isolated avian gut-specific gene
CC expression control region), expressing a heterologous polypeptide in a
CC host cell (by transfecting a eukaryotic cell with the recombinant DNA
CC molecule, and culturing the transfected cell in a medium suitable for
CC expression of a heterologous polypeptide under the control of an avian
CC intestinal fatty acid binding protein (IFABP) or cp35 gene expression
CC control region encoded by the recombinant DNA molecule), a eukaryotic
CC cell transformed with the expression vector (or its progeny, which
CC expresses a heterologous polypeptide) and a transgenic avian having a
CC heterologous polynucleotide sequence comprising the nucleic acid insert.
CC The nucleic acids are useful for regulating heterologous nucleic acids in
CC transgenic avians, as probes in nucleic acid hybridisation assays for
CC detecting the IFABP gene expression control region, and for generating
CC transgenic birds. The present sequence is the Chicken intestinal fatty
CC acid binding protein, IFABP, gene, 5' region.

XX S0 Sequence 2381 BP; 775 A; 428 C; 469 G; 709 T; 0 U; 0 Other;

Query Match 100.0%; Score 512; DB 11; Length 2381;

Best Local Similarity 100.0%; Pred. No. 3.3e-129; Mismatches 0; Indels 0; Gaps 0;

Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGTGGATCTTCATCTCATGCTTATTTATGTGAGTAGAATAGATTCTCACCATTGA 60
DB 1115 CAGTGGATCTTCATCTCATGCTTATTTATGTGAGTAGAATAGATTCTCACCATTGA 1174
QY 61 ATGACAAAGCAGAGATTGTTGTTTATCTGTGGTAAATACGTTTCTCCAGTTGAT 120
DB 1175 ATGACAAAGCAGAGATTGTTGTTTATCTGTGGTAAATACGTTTCTCCAGTTGAT 1234
QY 121 AAAGACCTCCACAGATATAAGCTATGCAACAAAGAAATGTCATATCATTCTCT 180
DB 1235 AAAGACCTCCACAGATATAAGCTATGCAACAAAGAAATGTCATATCATTCTCT 1294
QY 181 AGTCTCATTTATTTTCTATAGATAGCCGGTTTTCTCTCACTCAATTAAGTAGAC 240
DB 1295 AGTCTCATTTATTTTCTATAGATAGCCGGTTTTCTCTCACTCAATTAAGTAGAC 1354
QY 241 AGATGAATGGGTTATAGTACTGTTTATTAAGAAAGTAAATAGTACTATCATCTTTG 300
DB 1355 AGATGAATGGGTTATAGTACTGTTTATTAAGAAAGTAAATAGTACTATCATCTTTG 1414
QY 301 AGGCATTAAGGAGGAGAGATTCAGCAACAGTGTGCTTCAAGTGGAAACAAGTTAA 360
DB 1415 AGGCATTAAGGAGGAGAGATTCAGCAACAGTGTGCTTCAAGTGGAAACAAGTTAA 1474
QY 361 ACTAAAGTACCCCTCTCTTGACAAGTCAATGCAAGTTGAGCTTTAGCCAGCA 420
DB 1475 ACTAAAGTACCCCTCTCTTGACAAGTCAATGCAAGTTGAGCTTTAGCCAGCA 1534
QY 421 TCATCATGTAATATGCTTCTGTAATAGCCCTGTCATTAATCTCTTTGCAAACTCTGC 480
DB 1535 TCATCATGTAATATGCTTCTGTAATAGCCCTGTCATTAATCTCTTTGCAAACTCTGC 1594
QY 481 TACTTACCAGAAAGTCTGCTTACAGACAGAAAG 512
DB 1595 TACTTACCAGAAAGTCTGCTTACAGACAGAAAG 1626

RESULT 2

ADL90128
ID ADL90128 standard; DNA; 336 BP.

XX AC ADL90128;

XX DT 20-MAY-2004 (first entry)

XX DE Chicken intestinal fatty acid binding protein, IFABP, gene, promoter.

XX KM Chicken; ds; intestinal fatty acid binding protein; IFABP;

KM gut specific promoter; transgenic; promoter.

XX OS Gallus gallus.

XX FN US2003177516-A1.

XX PD 18-SEP-2003.

XX PF 14-MAR-2002; 2002US-00099663.

XX PR 14-MAR-2002; 2002US-00099663.

XX PA (HORS/) HORSEMAN N D.

XX PA (PRAT/) PRATT S L.

XX PI Horseman ND, Pratt SL;

XX DR WPI; 2003-898653/82.

XX PT New nucleic acid molecule comprising an isolated avian gut-specific gene
XX PT expression control region, useful for regulating heterologous nucleic
XX PT acids in transgenic avians, and for generating transgenic birds.

XX PS Claim 1; SEQ ID NO 2; 28bp; English.

XX CC The invention relates to an isolated nucleic acid comprising an isolated
XX CC avian gut-specific gene expression control region appearing as
XX CC ADL90127(chicken intestinal fatty acid binding protein, IFABP, gene, 5'
XX CC region or ADL90128 (chicken IFABP promoter) or its degenerate variant.
XX CC Also included are a recombinant DNA molecule comprising an isolate avian
XX CC gut-specific gene expression control region operably linked to a nucleic
XX CC acid insert encoding a polypeptide, an expression vector that integrates
XX CC into a host cell (and comprising the isolated avian gut-specific gene
XX CC expression control region), expressing a heterologous polypeptide in a
XX CC host cell (by transfecting a eukaryotic cell with the recombinant DNA
XX CC molecule, and culturing the transfected cell in a medium suitable for
XX CC expression of a heterologous polypeptide under the control of an avian
XX CC intestinal fatty acid binding protein (IFABP) or cp35 gene expression
XX CC control region encoded by the recombinant DNA molecule), a eukaryotic
XX CC cell transformed with the expression vector (or its progeny, which
XX CC expresses a heterologous polypeptide) and a transgenic avian having a
XX CC heterologous polynucleotide sequence comprising the nucleic acid insert.
XX CC The nucleic acids are useful for regulating heterologous nucleic acids in
XX CC transgenic avians, as probes in nucleic acid hybridisation assays for
XX CC detecting the IFABP gene expression control region, and for generating
XX CC transgenic birds. The present sequence is the Chicken intestinal fatty
XX CC acid binding protein, IFABP, gene, promoter.

XX S0 Sequence 336 BP; 116 A; 62 C; 63 G; 95 T; 0 U; 0 Other;

Query Match 63.7%; Score 326; DB 11; Length 336;

Best Local Similarity 100.0%; Pred. No. 9.3e-79; Mismatches 0; Indels 0; Gaps 0;

Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 ATTATATTTTCTATAGATAGCCGGTTTTTCTACCACTCAATAGATGAAGCAATG 246
DB 1 ATTATATTTTCTATAGATAGCCGGTTTTTCTACCACTCAATAGATGAAGCAATG 60
QY 247 AATGGTTAGTACCTGTTTATTAAGAAAGTAAATAGTACTATCATCTTTGAGGCA 306
DB 61 AATGGTTAGTACCTGTTTATTAAGAAAGTAAATAGTACTATCATCTTTGAGGCA 120
QY 307 TAAGGAGGAGAGATTTCAGCAACAGTGTGCTTACAGTGGAAACAAGTTAACTAAA 366
DB 121 TAAGGAGGAGAGATTTCAGCAACAGTGTGCTTACAGTGGAAACAAGTTAACTAAA 180
QY 367 GTGACCCCTCTCTTGACAAGATGCAAGTTCAGCTTTAGCCAGCCACATCATCA 426
DB 181 GTGACCCCTCTCTTGACAAGATGCAAGTTCAGCTTTAGCCAGCCACATCATCA 240
QY 427 TGTAAATGCTTCTCGATTAAGCTGTGTTAATATCTCTTTGCAAAAGTCTGCTACTTA 486
DB 241 TGTAAATGCTTCTCGATTAAGCTGTGTTAATATCTCTTTGCAAAAGTCTGCTACTTA 300

Oy	487	CCAGAACTCTGCTACAGACAGAAG	512
Dd	301	CCAGAACTCTGCTACAGACAGAAG	326
 RESULT 3 ACL35363			
ID	ACL35363	standard; cDNA; 2000 BP.	
XX	XX	ACL35363;	
DT	02-JUN-2005	(first entry)	
DE	Rice stress-regulated promoter SEQ ID NO:13926.		
KM	ss; abiotic stress tolerance; transgenic plant; plant; cereal;		
OS	agriculture.		
XX	Oryza sativa.		
PN	WO2003008540-A2.		
PD	30-JAN-2003.		
PF	21-JUN-2002; 2002WO-US019668.		
PR	22-AUG-2001; 2001US-0300113P.		
PR	24-AUG-2001; 2001US-0314662P.		
PR	26-SEP-2001; 2001US-0325277P.		
PR	21-NOV-2001; 2001US-0332132P.		
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.		
PI	Krepe J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,		
PI	Moughamer T, Provart N, Rieke D, Zhu T;		
XX	WPI; 2003-248011/24.		
DR	New stress-responsive nucleic acid, useful for altering the		
PT	responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold		
PT	stress, salt stress or osmotic stress.		
PS	Claim 48; SEQ ID NO 13926; 89pp; English.		
XX	The invention relates to novel abiotic stress responsive polynucleotides		
XX	and polypeptides. Also disclosed are vectors, expression cassettes, host		
CC	cells, and plants containing such polynucleotides. Also disclosed are		
CC	methods for using the polynucleotides and polypeptides to alter the		
CC	responsiveness of a plant to abiotic stress. The invention is useful in		
CC	agriculture. The nucleic acid is useful for determining whether a test		
CC	plant has been exposed to an abiotic stress condition. It is also useful		
CC	for selecting an agent that alters abiotic stress regulated		
CC	polynucleotide expression in a plant cell, and to identify a homolog or		
CC	ortholog to an abiotic stress responsive polynucleotide. The nucleic acid		
CC	molecule and the polypeptide encoded by it are useful in altering the		
CC	responsiveness of a plant to an abiotic stress, such as cold stress, salt		
CC	stress, osmotic stress or any of their combinations. The present sequence		
CC	is used in the exemplification of the invention		
SQ	Sequence 2000 BP; 540 A; 344 C; 416 G; 428 T; 0 U; 272 Other;		
Query Match	8.4%; Score 42.8; DB 11; Length 2000;		
Best Local Similarity	11.7%; Pred. No. 0.3;		
Matches	38; Conservative 148; Mismatches 140; Indels 0; Gaps 0		
Oy	14	ATCTCATGCTTATTATGTGGAGTAGAATGATTCACCAATAATTCAGATGCACAACCG	73
Dd	19	WKYMMKKMGKAYMKGAKWMWTMSAVYAMMYAKKKMTBARKCTCSCKYMSVA	78
Oy	74	AGATTGGTTTATCTGTTGGGTAAATAGCTTTTCTCAGTTGTATTAAGACCCTCCA	133
Dd	79	SGKHMMWSMARGSNKTYYTGGRMRKYAYIKKMTMTMAAWSTRTMTTKACTCYTT	138

OY 134 CGAGATTAAGGCCATGCAACAAAGAATACTCATCTTACTGTTCATTATTA 193
 DB 139 MYKMTMYTRMAAAMRWKKMKYKKWAKRGSSAKAYRKRKCCTYYANRRMANMNTTKC 198
 OY 194 TTTTCATTAGTAGCCGGTTTTTTTACTACAACCTCAAATPAAGTGAACAGATGAATGGT 253
 DB 199 TMSWVAAMWTCAATMAAWATAATTTTTRBAAWMTTBRAAMMTBRVAMMYTWNAKKMAVKYAT 258
 OY 254 TAGTGACGCGTTATTAAGAAGTAATAAAGTACTATCATCATTTTGAGCAATAAGGA 313
 DB 259 KMTTYAAMRWMMNNMWSRTTYRMMTWAMMYTWMSWYTWAAMSWMCMMTWSRRKMWR 318
 OY 314 GCGAGAGATTCAGCAAAACAGTGTCT 339
 DB 319 YKSRRRTTWTMTWTTWMAAMKKWKX 344

RESULT 4
 ADA71938/C
 ID ADA71938 standard; DNA; 2000 BP.
 AC ADA71938;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Rice gene, SEQ ID 5263.
 XX
 PL plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene; ds.
 XX
 OS *Oryza sativa*.
 XX
 PN WO2003000898-A1.
 XX
 PD 03-JAN-2003.
 PF 22-JUN-2001; 2001WO-IB001105.
 PR 22-JUN-2001; 2001WO-IB001105.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
 PI Katagiri F, Qian S, Tao Y, Whitlham S, Xie Z, Zhu T, Zou G;
 DR WPL: 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX
 PS Claim 27; SEQ ID NO 5263; 899pp; English.

CC The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX
 SO Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other:

Query Match 8.3%; Score 42.6; DB 8; Length 2000;
 Best Local Similarity 10.6%; Pred.No. 0.34;
 Matches 44; Conservative 194; Mismatches 170; Indels 1; Gaps 1.

1 CAGTGCATCCCTCATCTCAAGCTTATTAATGAGAGTAGAATGATTCACCAATTAAGA 60

```

Db      1016 MRRYTRSMSTWYAMMKYTKYMTAYSSTWKMYAYKRAYAMRSRKTMTGCKMAT 957
Qy      61  ATGCAAGAGAGAGATTTGCTTTATCTGTTGGTAATAAGCTTTTCTCCAGTTGAT 120
Db      956 YCGTMMAGKMRMAMCMYCCMMKMKMTSCMMYKMYRWSCMYTMMGAMRAYAM 897
Qy      121  AAAGACCTCCACAGATTAAGTCTATGCAACAAGAA- AATGCAATATCTTCT 179
Db      896 RRRRTYKMSWRMYMTWTMAMTMCMCAKMYATGATMMMYTTCYATCA 837
Qy      180 TAGCTCATTTATTTTCTATTAATAGCCGTTTCTTACTCAACTCAATAAGATGA 239
Db      836 KCKTKMAMTMMTWTACAMRATSWMRAMKRYMRAAAYMWRMWRCKMGNARMK 777
Qy      240 CAGATGATGAGTGAAGTCTTATTAAGAAGATTAATAAGATCATCATATTT 299
Db      776 SRYMKMKKATATRYMMAMTMMWRMKSTRMMSGMRMWSMRYSRKKACTRY 717
Qy      300 GAGCCATAAGGAGGAGAGATTCAGCAAAAGTGTCTTACAGATGGAAGAAAGATTA 359
Db      716 ASSRMTKRAKRSYRRMYRMYRKGTWYRYRYSRSCMTRAMSRKRRMAGSMKSCM 657
Qy      360 AACTAAAGTACCCCTCTGCAAGATCATGCCACAGTGAAGCTTTAGCCAGC 416
Db      656 YMRGASMWYSKXSCAKCKKTRMTSTMTGMTSYKSMSTSKMSYMGKM 600

```

RESULT 5
AAS45387/c
ID AAS45387 standard; DNA; 5908 BP.

AC AAS45387;

DT 18-DEC-2001 (first entry)

DE Chemically pretreated complementary DNA associated with cell cycle #46.

XX Cell cycle; human; CpG dinucleotide; cytosine methylation; HIV; aging;
 XX human immunodeficiency virus; neurodegenerative disorder; solid tumour;
 XX graft-versus-host disease; glomerular disease; Lewy body disease; cancer;
 XX arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;
 XX immunosuppressive; antitumour; cytosolic; antiarteriosclerotic; ds;
 XX PCR primer.

OS Homo sapiens.

PN WO200168911-A2.

PD 20-SEP-2001.

PF 15-MAR-2001; 2001WO-EP002945.

PR 15-MAR-2000; 2000DE-01013847.

PR 06-APR-2000; 2000DE-01019058.

PR 07-APR-2000; 2000DE-01019173.

PR 30-JUN-2000; 2000DE-01032529.

PR 01-SEP-2000; 2000DE-01043826.

PA (EPIG-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

PT WPI; 2001-602751/68.

XX Designing primers and probes for analyzing diseases associated with

XX cytosine methylation state e.g. arthritis, cancer, aging,

XX arteriosclerosis comprising fragments of chemically modified genes

XX associated with cell cycle.

XX Claim 1; SEQ ID NO 92; 28bp; English.

XX Sequences AAS45296-AAS45520 represent chemically pretreated genomic DNA

XX molecules associated with the cell cycle and specific PCR primers of the

CC invention. The sequences are useful for detecting the methylation state
 CC of all CpG dinucleotides in a sequence and therefore for analysing
 CC associated diseases. By analysing cytosine methylations in the pretreated
 CC DNA, genetic and/or epigenetic parameters for the diagnosis and therapy
 CC of existing diseases or the predisposition to specific diseases can be
 CC ascertained. The parameters may be compared to another set of genetic
 CC and/or epigenetic parameters, the differences serving as basis for
 CC diagnosis and/or prognosis events which are disadvantageous to patients.
 CC The sequences of the invention are useful for the diagnosis and therapy
 CC of HIV infection, neurodegenerative disorders, graft-versus-host disease,
 CC arteriosclerosis, solid tumours and cancers

SQ Sequence 5908 BP; 1706 A; 149 C; 1421 G; 2632 T; 0 U; 0 Other;

Query Match 8.3%; Score 42.4; DB 4; Length 5908;

Best Local Similarity 54.5%; Pred. No. 0.55; 71; Indels 0; Gaps 0;

Matches 85; Conservative 0; Mismatches 0

```

Qy      138 TATTAAGTCTTATGCAAGAAAGAAATGTCATATCATTTCTTATCTATTTATTT 197
Db      4703 TATTAAGTCTTATTAATAAACAATAATATATATCTATTAACCACTTTTAA 4644
Qy      198 CATTAAGTACCGGTTTCTTACTCAACTCAATAAGATGAACAGATGAGGTTAGT 257
Db      4643 TATTAATTTTATTTTCACTACCTTCTCTTAATAATTAATAATTAATCTATACCTTAT 4584
Qy      258 GACTGTTTATTAAGAAGATTAATAAGATCATATCA 293
Db      4583 TCCCTTTATCCAAATTAATTAATCTATATACGAACA 4548

```

RESULT 6
ABK28232/c
ID ABK28232 standard; DNA; 5908 BP.

AC ABK28232;

DT 23-APR-2002 (first entry)

DE DNA transcription associated complementary genomic DNA #53.

XX DNA transcription associated gene; peptide nucleic acid; pNA-oligomer;
 XX pNA; cytosine methylation state; SNP; retroviral infection; gene; ds;
 XX single nucleotide polymorphism; adenoviral deamase deficiency; cancer;
 XX viral infection; Sezary syndrome; haematological disorder; tuberculosis;
 XX immunological disorder; Werner syndrome; developmental disorder;
 XX psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;
 XX neurodegenerative disorder; Maardenburg syndrome; Niemann-Pick disease;
 XX myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;
 XX angiodysgenesis; congenital heart disease; HDR syndrome; gene therapy;
 XX polyglutamine disorder; solid tumour.

OS Unidentified.

PN WO200192565-A2.

PD 06-DEC-2001.

PF 06-APR-2001; 2001WO-EP003973.

PR 06-APR-2000; 2000DE-01019058.

PR 07-APR-2000; 2000DE-01019173.

PR 30-JUN-2000; 2000DE-01032529.

PR 01-SEP-2000; 2000DE-01043826.

PA (EPIG-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

PT WPI; 2002-090046/12.

XX New nucleic acids or oligomers, useful for diagnosing or treating

XX

OS Homo sapiens.
XX WO200177375-A2.
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-EP003968.
XX
XX 06-APR-2000; 2000DE-01019058.
PR 07-APR-2000; 2000DE-01019173.
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI, 2002-017470/02.
XX
XX New nucleic acid sequences from chemically modified genes associated with
PT gene regulation, useful for analyzing cytosine methylations for diagnosis
PT and therapy of diseases e.g. severe combined immunodeficiency disease.
XX
XX Claim 1; SEQ ID NO 211; 26bp; English.
XX
XX The invention relates to 224 nucleic acid sequences comprising at least
CC 18 bases of a chemically pretreated gene associated with gene regulation
CC selected from 43 known genes (or complementary sequences). The chemical
CC pretreatment converts cytosine bases unmethylated at the 5-position to
CC uracil or another base with hybridisation behaviour dissimilar to
CC cytosine, to enable analysis of cytosine methylations. The DNA sequences,
CC oligomers (or sets/arrays) and method are useful in the diagnosis of
CC diseases (or predisposition to diseases) associated with gene regulation
CC and in therapy of such diseases, by enabling analysis of the cytosine
CC methylation patterns of such genes, kits are provided. They are
CC especially useful in diagnosis and therapy of e.g. severe combined
CC immunodeficiency disease, cardiac disorders, haemophilia, solid tumours
CC and cancer, Werner syndrome, asthma, HDR syndrome, Saethre-Chotzen
CC syndrome, renal disease, presenile dementia, graft versus-host disease. The
CC present sequence is a sequence included in the sequence data for this
CC specification and is associated with the human gene regulation-associated
CC genes. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_sequences
XX
XX Sequence 10048 BP; 2873 A; 57 C; 1888 G; 5225 T; 0 U; 5 Other;
SQ
Query Match 8.2%; Score 41.8; DB 6; Length 10048;
Best Local Similarity 46.4%; Pred. No. 0.95;
Matches 136; Conservative 0; Mismatches 157; Indels 0; Gaps 0;
OY 100 ATACGTTTCTCCAGTTGTAATAAGACCTCCACCACTATTAAGTCTATGCAACAAG 159
DB 5471 ATATTTATTTATCTTAAACCTAAATTCCTCTCACAAAAAAAATTAATCTATAC 5412
OY 150 AAAATGTCATACATCTCTTCTAGTCTCATTTATTTTCTATAGTAGCGGTTTTTAC 219
DB 5411 TCCAAATCCAAAAATCTATTTTAAACCTATTAACCTCTCAACAATATCTTTTAC 5352
OY 220 TACAACTCAATTAAGATGAACGAATGAATGAGTTAGTACTGTTTAAAGAGAGTAA 279
DB 5351 AAATTACTTAATTAATAAACACACATTAATTCCTTAATCTATTAATCAAAATACAAAAAAC 5292
OY 280 TAAAGATACATCATTTGAGGCAATTAAGGAGGAGAGATTCACCAACAAGTGTGCT 339
DB 5291 TATTAATTAATTTATTTAAATTAATTTCACTTAATTAATTTTAAATTAATTAATTTTCA 5232
OY 340 TACAAAGTGAACAAGTAAAGTAAAGTGAACCCCTCTCTGCAAGATCA 392
DB 5231 AAACAAAACAAAACAAAACAAAACAAAACCAATTTTCAAACTATATCTA 5179

5 RESULT 9

AA546439/c
ID AA546439 standard; DNA; 7040 BP.
XX
XX AC AA546439;
XX
XX 18-DEC-2001 (first entry)
XX
XX Tumour suppressor gene derived chemically modified sequence #161.
XX Human; tumour suppressor gene; oncogene; antitumour; cytosine; cancer;
XX tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
XX cytosine methylation; ds.
XX
XX Homo sapiens.
XX
XX WO200168912-A2.
XX
XX 20-SEP-2001.
XX
XX 15-MAR-2001; 2001WO-EP002955.
XX
XX 15-MAR-2000; 2000DE-01013847.
PR 06-APR-2000; 2000DE-01019058.
PR 07-APR-2000; 2000DE-01019173.
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI, 2001-602752/68.
XX
XX Fragments of chemically modified genes associated with tumor suppressor
PT genes and oncogenes, useful in designing primers and probes for analyzing
PT diseases associated with cytosine methylation state e.g. cancer.
XX
XX Claim 1; SEQ ID NO 161; 27bp; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of 18
XX bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
XX bisulphite, of genes associated with tumor suppression and oncogenes
XX having a sequence taken from 536 (actually 533 since numbers 408, 458 and
XX 500 are missing from the sequence listing) sequences (Ss) and sequences
XX complementary to (Ss). The nucleic acid may be a peptide nucleic acid-
XX oligomer (PNA) of at least 9 nucleotides and may form part of a set of
XX probes for detecting the cytosine methylation state and/or single
XX nucleotide polymorphisms and also to be used in an array for analysing
XX probes can also be used in a method for ascertaining genetic and/or
XX epigenetic parameters for the diagnosis and/or therapy of existing
XX diseases or the predisposition to specific diseases, by analysing
XX cytosine methylations. The parameters may be compared to another set of
XX genetic and/or epigenetic parameters, the differences serving as basis
XX for diagnosis and/or prognosis events which are disadvantageous to
XX patients. The present sequence is one of the 533 genomic sequences
XX derived from tumor suppressor genes and oncogenes. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_sequences
SQ
Sequence 7040 BP; 2018 A; 184 C; 1566 G; 3272 T; 0 U; 0 Other;
Query Match 7.9%; Score 40.6; DB 4; Length 7040;
Best Local Similarity 50.8%; Pred. No. 1.8;
Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
OY 101 TACGTTTCTCCAGTTGTAATAAGACCTCCACCACTATTAAGTCTATGCAACAAG 160
DB 3586 TATATTTTAAAAATCAAAATTAATTCCTCTCTTAACACCTATATTAATTT 3527
OY 161 AAATGTCATACATCTCTAGTCTCATTTATTTTCTATTAATTAAGCGGTTTTTACT 220

Db 3536 ATATATTAACGACATTTTAAATTAACAACTCTATTTTAAATATATTAACCTTAAAAA 3467
Oy 221 ACAACTCAAAATAGATGAGACAGATGATGGTTAGTGCCTGTTTAAAGAGACTAAT 280
Db 3466 ATTAATTTAATAACCTTAACCAATTTTCAAAAAACAATTTAAATCTATTAACCTTAAT 3407
Oy 281 AAAAGTACTAT 291
Db 3406 AAAATTTCTAT 3396
RESULT 10
ABK33963/c
ID ABK33963 standard; DNA; 7040 BP.
XX
AC ABK33963;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human DNA for staging of Astrocytomas #24.
XX
KM Human; ds; astrocytoma; cytosstatic; staging; cysteine methylation; CpG;
KM bisulphite; brain tissue; MALDI; ESI; electron spray mass spectrometry;
KM matrix assisted laser desorption/ionization mass spectrometry.
XX
OS Homo sapiens.
XX
PN WO200202808-A2.
XX
PD 10-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007538.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-171649/22.
XX
PT Novel chemically modified genomic DNA sequences, useful in the
PT characterization, classification, differentiation, grading, staging,
PT treatment and/or diagnosis of astrocytomas or predisposition to
PT astrocytomas.
XX
PS Claim 1; SEQ ID NO 47; 37pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence (I) of at
CC least 18 bases in length of a segment of chemically pre-treated genomic
CC DNA which has any one of the sequences of (ABK33915-ABK34032) or its
CC complement. Also included are an oligonucleotide or peptide nucleic acid
CC (or set thereof) of at least 9 nucleotides which hybridises to (I),
CC primers for (I), probes for detecting cytosine methylation or single-
CC nucleotide polymorphisms (SNP) in (I), an array of oligomers or peptide
CC nucleic acids for analysing diseases associated with the methylation
CC states of the CpG dinucleotides of (I). The array is useful for
CC determining genetic and/or epigenetic parameters, classification,
CC differentiation, grading, staging, treatment and/or diagnosis of
CC astrocytomas, or the predisposition to astrocytomas by analysing cytosine
CC methylations, involves obtaining a biological sample containing genomic
CC DNA, extracting the genomic DNA, converting cytosine bases which are
CC unmethylated at the 5-position, in the genomic DNA sample, to uracil or
CC another base which is dissimilar to cytosine in terms of hybridisation
CC behaviour, by chemical treatment and amplifying chemically pre-treated
CC genomic DNA fragments using the array and a polymerase, where the
CC amplicates carry a detectable label. The method further involves
CC identifying methylation status of one or more cytosine positions, and
CC analysing methylation status of the cytosine positions by reference to
CC one or more data sets. The genomic DNA is chemically treated by using a
CC bisulphite, hydrogen sulphite or disulphite. The amplification step
CC amplifies DNA which is of particular interest in astrocytoma or brain

CC tissue, based on the specific genomic methylation status of brain
CC tissues, as opposed to background DNA. The amplicates carry a
CC fluorescent label or radionuclide. Optionally, the labels of the
CC amplicates are detachable molecule fragments having a typical mass
CC which are detected in a mass spectrometer. The fragments of chemically
CC pre-treated genomic DNA to be amplified, have a single positive or
CC negative charge for a better detectability in the mass spectrometer.
CC Preferably, the amplicates or fragments of the amplicates are
CC detected by matrix assisted laser desorption/ionization mass spectrometry
CC (MALDI) or using electron spray mass spectrometry (ESI). The present
CC sequence is one of the chemically pre-treated reference DNA samples of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 7040 BP; 2018 A; 184 C; 1566 G; 3272 T; 0 U; 0 Other:
Query Match 7.9%; Score 40.6; DB 6; Length 7040;
Best Local Similarity 50.8%; Pred. No. 1.8;
Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
Oy 101 TACGTTTCTCAGTTGATTAAGACCTCCACAGTATTAAGCTTATGCAACAANA 160
Db 3586 TATATTTAAAAAATCAATATTAATAATTCCTTTCTAAAAACACTATATATTAAT 3527
Oy 161 AAATGCAATACATCTCTAGTCTCATATATATTTTCATAGATAGCCGTTTTTACT 220
Db 3526 ATATATTAACCACTTTTATTAATTAACATTTCTTTTAAATATTAATTTAAAAA 3467
Oy 221 ACAACTCAAAATAGATGAGACAGATGATGGTTAGTGCCTGTTTAAAGAGACTAAT 280
Db 3466 ATTAATTTAATAACCTTAACCAATTTTCAAAAAACAATTTAAATCTATTAACCTTAAT 3407
Oy 281 AAAAGTACTAT 291
Db 3406 AAAATTTCTAT 3396
RESULT 11
AB210179/c
ID AB210179 standard; DNA; 7040 BP.
XX
AC AB210179;
XX
DT 16-JAN-2003 (first entry)
XX
DE Haematopoietic cell proliferation disorder related DNA sequence #319.
XX
KM Human; haematopoietic cell proliferation disorder; cytosstatic;
KM gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KM cytosine methylation state; gene; ds.
XX
OS Homo sapiens.
XX
PN WO20027272-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-EP003401.
XX
PR 26-MAR-2001; 2001US-0278333P.
XX
PA (EPIG-) EPIGENOMICS AG.
PI Berlin K, Braun A, Distler J, Guectig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Leesche R, Leu B;
PI Lewin A, Lipscher B, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI Schwoge I, Ziebrich H;
XX
DR WPI; 2003-018942/01.
XX
PT Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that

PT distinguishing between methylated and non-methylated CpG dinucleotides.
 XX
 XX Claim 28; SEQ ID NO 319; 117bp; English.
 PS
 CC The present invention describes a method for detecting and
 CC differentiating between haematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. AB209861 to AB211118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used: for
 CC differentiating between healthy haematopoietic cells and proliferative
 CC disorder haematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of haematopoietic cell proliferation disorder related DNA
 CC sequences. The nucleotide sequences from the present invention can also
 CC be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC haematopoietic cell proliferative disorders. The present method enables a
 CC highly specific classification of haematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients
 CC
 XX
 SQ Sequence 7040 BP; 2018 A; 0 C; 1566 G; 3456 T; 0 U; 0 Other;
 XX
 Query Match 7.9%; Score 40.6; DB 8; Length 7040;
 Best Local Similarity 50.8%; Pred. No. 1.8;
 Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
 QY 101 TACGTTTCTCCAGTGTATTAAGACCTCCACAGTATTAAGTCTATGCAACAAAGA 160
 Db TATATTTTAAAAAATCAATATTAATTAATCCCTTCTTCAAAAACCTTATATTAATTT 3527
 QY 161 AAATGTCATACATCTCTTACTCTCATTTATTTTCAATAGATGCGGTTTTTACT 220
 Db ATATATTAACCACTTTTATTAATTAACATTTTATTAATTAATTAATTAATTAATTA 3467
 QY 221 ACAACTCAATTAAGATGAACAGATGAATGGTGTAGTCACTGTTTATTAAGAGATAT 280
 Db ATTAATTAATTAACATTAACATTTTACAAACATTTTAAATCTATTAACCCCTTAAT 3407
 QY 281 AAAGATACCTAT 291
 Db 3406 AAAATTTCTAT 3396
 RESULT 12
 AB210033/C
 ID AB210033 standard; DNA; 7040 BP.
 XX
 XX AB210033;
 AC
 XX 16-JAN-2003 (first entry)
 DT
 XX
 XX Haematopoietic cell proliferation disorder related DNA sequence #173.
 DE
 XX
 XX Human; haematopoietic cell proliferation disorder; cytostatic;
 KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
 KM cytosine methylation state; gene; ds.
 XX
 XX Homo sapiens.
 OS
 XX
 XX MO20027272-A2.
 PN
 XX
 XX 03-OCT-2002.
 PD
 XX
 XX 26-MAR-2002; 2002WO-EP003401.
 PF
 XX
 XX 26-MAR-2001; 2001US-027633P.
 PR

XX
 PA (EPIC-) EPIGENOMICS AG.
 XX
 XX Berlin K, Braun A, Distler J, Guecig D, Howe A, Mueller J;
 PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Leech R, Leu E, Pellet C;
 PI Lewin A, Lipscher B, Mater S, Model F, Mueller V, Otto T, Pelet C;
 PI Schwope I, Ziebarth H;
 DR WPI; 2003-018942/01.
 XX
 XX
 PT Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent that
 PT distinguishes between methylated and non-methylated CpG dinucleotides.
 PS
 CC Claim 28; SEQ ID NO 173; 117bp; English.
 XX
 XX
 CC The present invention describes a method for detecting and
 CC differentiating between haematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. AB209861 to AB211118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used: for
 CC differentiating between healthy haematopoietic cells and proliferative
 CC disorder haematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of haematopoietic cell proliferation disorder related DNA
 CC sequences. The nucleotide sequences from the present invention can also
 CC be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC haematopoietic cell proliferative disorders. The present method enables a
 CC highly specific classification of haematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients
 CC
 XX
 SQ Sequence 7040 BP; 2018 A; 184 C; 1566 G; 3272 T; 0 U; 0 Other;
 XX
 Query Match 7.9%; Score 40.6; DB 8; Length 7040;
 Best Local Similarity 50.8%; Pred. No. 1.8;
 Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
 QY 101 TACGTTTCTCCAGTGTATTAAGACCTCCACAGTATTAAGTCTATGCAACAAAGA 160
 Db TATATTTTAAAAAATCAATATTAATTAATCCCTTCTTAAACCTTATATTAATTT 3527
 QY 161 AAATGTCATACATCTCTTACTCTCATTTATTTTCAATAGATGCGGTTTTTACT 220
 Db ATATATTAACCACTTTTATTAATTAACATTTTATTAATTAATTAATTAATTAATTA 3467
 QY 221 ACAACTCAATTAAGATGAACAGATGAATGGTGTAGTCACTGTTTATTAAGAGATAT 280
 Db ATTAATTAATTAACATTAACATTTTACAAACATTTTAAATCTATTAACCCCTTAAT 3407
 QY 281 AAAGATACCTAT 291
 Db 3406 AAAATTTCTAT 3396
 RESULT 13
 ADA20348/C
 ID ADA20348 standard; DNA; 7040 BP.
 XX
 XX ADA20348;
 AC
 XX
 XX 20-NOV-2003 (first entry)
 DT
 XX
 XX Prostate tumour related genomic DNA sample #7.
 DE
 XX
 XX cytostatic; gene therapy; genetic marker; epigenetic parameter;
 KW

XX	classificati	differentiat	diagnosis	prostate tumour	
KV	prostate cancer	cytosine methylation	uracil		
KM	single nucleotide polymorphism	SNP	prostate carcinoma	88.	
XX					
OS	Homo sapiens.				
XX					
PN	MO2002103042-A2.				
XX					
PD	27-DEC-2002.				
XX					
PF	14-JUN-2002; 2002WO-EP006605.				
XX					
PR	14-JUN-2001; 2001DE-01028508.				
XX					
PA	(EPIC-) EPIGENOMICS AG.				
XX					
PI	Dietler J, Model F, Adorjan P;				
XX					
DR	WPI, 2003-167536/16.				
XX					
PT	Determining genetic and/or epigenetic parameters, useful for the				
XX	classification, differentiation and/or diagnosis of prostate tumours or a				
PT	predisposition to prostate cancer, comprises analyzing cytosine				
XX	methylation.				
XX					
PS	Claim 28; Page 68-72; 376pp; English.				
XX					
CC	The invention relates to a method of determining genetic and/or				
CC	epigenetic parameters for the classification, differentiation and/or				
CC	diagnosis of prostate tumours or the predisposition to prostate cancer,				
CC	by analysing cytosine methylation in a sample of genomic DNA. The method				
CC	comprises chemically treating unmethylated cytosine bases at the 5-				
CC	position to uracil or another base, which is dissimilar to cytosine in				
CC	terms of hybridization behaviour; followed by amplifying at least one				
CC	fragment of the chemically pre-treated genomic DNA using sets of primer				
CC	oligonucleotides and a polymerase. The oligomers or probes derived from				
CC	them are useful for detecting the methylation state of all CpG				
CC	dinucleotides and/or single nucleotide polymorphisms (SNPs) in a				
CC	chemically pre-treated genomic DNA. They are all useful for treating				
CC	prostate carcinoma. This sequence represents a fragment of genomic DNA				
CC	used in the method of the invention.				
XX					
XX	Sequence 7040 BP; 2018 A; 184 C; 1566 G; 3272 T; 0 U; 0 Other;				
SO					
	Query Match	7.9%;	Score 40.6;	DB 8;	Length 7040;
	Best Local Similarity	50.8%;	Pred. No.1.8;		
	Matches	97;	Conservative	0;	Mismatches 94; Indels 0; Gaps 0;
QY	101 TACGTTTCTCCAGTGTATATAAAGACCCCTCCACGATATAAAGTCTATGCAACAAGA	160			
DB	3586 TATATTTTAAAAAACAATATAATAAATTCCTTCTGTAAAAACATTATATATTAATT	3522			
QY	161 AAAGTCAATACATCTCTCTAGCTCANTATATTTTCATTAGATAGCCGGTTTTTACT	220			
DB	3526 ATATAATAACCACTTTTATATAATAACATTCATATTTTAAAAATATATTAACCTTTAAAAA	3467			
QY	221 ACAACTCAATTAAGATGAACGAATGAATGSGTGTAGTACGTATTATAAAGAAGTAAAT	280			
DB	3466 ATTAATTTATATAACATTAACCACTTTACAAAAACAATTTAAATCTATTAATACCTTAAT	3407			
QY	281 AAAGATACTAT 291				
DB	3406 AAAATTTCTAT 3396				
XX					
XX	RESULT 14				
XX	ADA84155/C				
XX	ID ADA84155 standard; DNA; 7040 BP.				
XX	ADA84155;				
XX					
XX	20-NOV-2003 (first entry)				
XX					

DE		Human renal/prostate carcinoma associated DNA SEQ ID NO:13.
XX		
KM	d8;	renal cancer; prostate cancer; cytosine methylation; single nucleotide polymorphism; histological; cytological.
XX		
OS	Homo sapiens.	
PN	WO2002103041-A2.	
XX		
PR	14-JUN-2001; 2001DE-01028509.	
XX		
DR	14-JUN-2002; 2002MO-BP066603.	
PD	27-DEC-2002.	
PI	Dierler J, Model F, Adorjan P;	
PA	(EPIG-) EPIGENOMICS AG.	
XX		
WPI	2003-183991/18.	
PT		
PT		
PT		
PS	Claim 1; Page 50-51; 21pp; English.	
CC	The invention relates to a novel method for characterizing, classifying and/or differentiating renal and prostate cancer. The method comprises extracting genomic DNA from a biological sample, converting cytosine bases (by chemical treatment) that are unmethylated at the 5-position to uracil or another base, and amplifying at least one fragment of the chemically pretreated genomic DNA using sets of primer oligonucleotides and a polymerase. The method is useful for detecting the cytosine methylation state and/or single nucleotide polymorphisms in genomic DNA, particularly for characterizing, classifying and/or differentiating renal and prostate cancers. The oligomers are useful as primer oligonucleotides for the amplification of any of the 112 DNA sequences of the invention. The set of oligomer probes is useful for detecting the cytosine methylation state and/or single nucleotide polymorphisms in any of the 112 chemically pretreated genomic DNA sequences. The method is also useful for identifying the tissue of origin of cancer cells. The method allows the classification, differentiation and/or diagnosis of cancer tissues using minute samples which would be inadequate for histological or cytological analysis. The present sequence represents one of the 112 DNA sequences of the invention.	
SQ	Sequence 7040 BP; 2018 A; 184 C; 1566 G; 3272 T; 0 U; 0 Other;	
Query Match	7.9%; Score 40.6; DB 8; Length 7040;	
Best Local Similarity	50.8%; Pred. No. 1.8;	
Matches	97; Conservative 0; Mismatches 94; Indels 0; Gaps 0	
Gy	101 TAGCTTTTCCTCAAGTGTATTAAGAACCCTCCCACGATATAAAGTCTATGCACAAGAA 160	
Db	3586 TATATTTTTAAAAAAAATCAATATATAAATTCCCTTCTTAABAACCTATATATTAAT 3522	
Gy	161 AAATGTCATATCATCTCTTAGTGCTCATTATATATTTTCATATAGATGCCGGTTTTACT 220	
Db	3526 ATATATATACCAATTTATATATATTAACATCTTATTTTTTAAATATATTAATCAATTA 3467	
Gy	221 ACAACTCAATATAGATGAAGAACGANTGCGATGCGATGTTATATAAGAGATTAAT 280	
Db	3466 ATTAATTTATATTAACATTAACATTTACAAAAACAATTTAAAACTATATTAACCCCTAAAT 3407	
Gy	281 AAAAGATACTAT 291	
Db	3406 AAAATTTCTAT 3396	
RESULT 15		
ADBE84191/C		

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 28, 2006, 05:53:57 ; Search time 2625.13 Seconds
(without alignments)
10906.400 Million cell updates/sec

Title: US-10-099-663-1_COPY_1115_1626

Perfect score: 512
Sequence: 1 cagtgatccttcacatcatc.....gtctgcctacagacagaag 512

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 2795965780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: gb_eat1:*
2: gb_eat3:*
3: gb_eat4:*
4: gb_eat5:*
5: gb_eat6:*
6: gb_hc:*
7: gb_eat2:*
8: gb_eat7:*
9: gb_eat8:*
10: gb_eat9:*
11: gb_gsa1:*
12: gb_gsa2:*
13: gb_gsa3:*
14: gb_gsa4:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47.6	9.3	1206	13	CL641444 CH213-11N
2	46	9.0	987	14	AL104456 Drosophila
3	45.2	8.8	1101	14	CNS000E08
4	44.4	8.7	910	11	AZ541558 ENTB15TF
5	44	8.6	894	11	BH155022 ENT135TF
6	44	8.6	1200	14	CNS0016CI Drosophila
7	43.8	8.6	1101	14	CNS0039G Drosophila
8	43.6	8.5	1158	12	CL108838
9	43.6	8.5	927	12	CG769354 TcB41.1_H
10	43.2	8.4	443	1	AV008603
11	43.2	8.4	784	14	DX034017 KBFB031L0
12	43.2	8.4	886	14	DX077691 KBFB089F2
13	43	8.4	523	12	CE169553 CIGF-g88-
14	43	8.4	811	13	DU115753 KBFH111N1
15	43	8.4	879	11	BH137539 ENTNA83TR
16	43	8.4	943	11	BH147340 ENTOM34TR
17	42.6	8.3	753	10	DM270229 UT-S-GS0-
18	42.6	8.3	1101	14	CNS0026Z
19	42.4	8.3	500	1	AU088319

20	42.4	8.3	1101	14	CNS0021J	AL061936 Drosophila
21	42.2	8.2	584	8	CO281770	CO281770 ERI58751.
22	42.2	8.2	851	11	AZ680935	AZ680935 ENTQ31TF
23	42.2	8.2	877	11	BH166323	BH166323 ENTRX53TR
24	42.2	8.2	988	12	CG625604	CG625604 CGUK31TV
25	42.2	8.2	1137	1	AJ926577	AJ926577 AJ926577
26	42	8.2	494	9	CX352331	CX352331 cethchmho
27	42	8.2	745	9	CX786928	CX786928 HES3C_66
28	42	8.2	1342	13	CL504041	CL504041 SAIL_733
29	41.8	8.2	690	13	CL756493	CL756493 OR_BBA012
30	41.8	8.2	890	11	BH135160	BH135160 ENT0B35TF
31	41.6	8.1	485	5	CK62482	CK62482 LE22737.3
32	41.6	8.1	546	3	BW516697	BW516697 BW516697
33	41.6	8.1	560	4	BW541385	BW541385 BW541385
34	41.6	8.1	584	3	BW518866	BW518866 BW518866
35	41.6	8.1	670	11	BZ087462	BZ087462 1K30b01.
36	41.6	8.1	1101	14	CNS0145U	AL103740 Drosophila
37	41.6	8.1	1173	12	CC215257	CC215257 CH261-117
38	41.4	8.1	444	3	BW139785	BW139785 BW139785
39	41.4	8.1	504	3	BW269615	BW269615 BW269615
40	41.4	8.1	574	1	AV680413	AV680413 AV680413
41	41.4	8.1	720	3	BW135968	BW135968 BW135968
42	41.4	8.1	725	3	BW186283	BW186283 BW186283
43	41.4	8.1	730	3	BW140184	BW140184 BW140184
44	41.4	8.1	732	3	BW127139	BW127139 BW127139
45	41.4	8.1	759	7	AV877102	AV877102 AV877102

ALIGNMENTS

RESULT 1
CL641444/C
DEFINITION CH213-11N06.SP6 CH213 Gasterosteus aculeatus genomic clone
CH213-11N06 3', genomic survey sequence.
CL641444
CL641444.1 GI:49660868

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
GSS.
Gasterosteus aculeatus (three spined stickleback)
Gasterosteus aculeatus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Gasterosteiformes;
Gasterosteidae; Gasterosteus.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 1206)
Kingsley,D., Grimwood,J., Dickson,M., Schmutz,J. and Myers,R.M.
Expressed sequence tags from Gasterosteus aculeatus
Unpublished (2004)
Contact: Grimwood, Jane
Stanford Human Genome Center
Stanford University School of Medicine
975 S California Avenue, Palo Alto, CA 94304, USA
Tel: 650 320 5917
Fax: 650 320 5801
Email: jane@hgsc.stanford.edu

FEATURES
source
Class: BAC ends
High quality sequence start: 16
High quality sequence stop: 503.
Location/Qualifiers
1..1206
/organism="Gasterosteus aculeatus"
/mol_type="genomic DNA"
/strain="Salmon River"
/db_xref="taxon:69293"
/clone="CH213-11N06"
/sex="Mixed"
/cell_type="Blood"
/clone_11b="CH213"
/note="Vector: pTRABAC2.1, Site 1: EcoRI. The sequence of the clone was established as a mapping and sequencing collaboration at the Stanford Genome Evolution Center,

funded by the NIH Centers of Excellence in Genomic Science (CEGS) Initiative (<http://cegs.stanford.edu>). The clone was isolated from the BAC library CHORI-213 built by Pictet de Jonckheere in collaboration with the Stanford Genome Evolution Center (<http://www.chori.org/bacpac/>). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering_information.htm).

ORIGIN

Query Match 9.3%; Score 47.6; DB 13; Length 1206;
Best Local Similarity 45.3%; Pred. No. 0.23;
Matches 170; Conservative 0; Mismatches 204; Indels 0; Gaps 0;

QY 11 TTCACTCATGCTTATATGAGTGAATGATTCACCAATTAAGATGACAAAG 70
DB 999 TTTACTATATATTTTATATGAAATCATATCTAATTAACATTAAGATTAAG 940
QY 71 CAGGATTTGTTTATGCTGTTGGTAAATACGTTTCTCCAGTTGATTAAGACCTC 130
DB 939 CTATATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 880
QY 131 CCACAGATTAATTAATGATGCAACAAAGAAATGCAATGCTCTTATAGTCAATTA 190
DB 879 TAAATATATATATATATATCAATTAATTAATTAATTAATTAATTAATTAAT 820
QY 191 TTATTTTCAATTAATGACCGGTTTATTAATCAATCAATTAATTAATTAATTA 250
DB 819 ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 760
QY 251 GGTATGACGCTTTTAAAGAGATTAATTAATTAATTAATTAATTAATTAATTA 310
DB 759 TATATGAATTTCCATTAACAAATTAATTAATTAATTAATTAATTAATTAATTA 700
QY 311 GAGGAGAGATTCAGCAACAGTGTCTTACAACTGAAAGCAATTAATTAATTA 370
DB 699 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 640
QY 371 CCCCCCTCTTGC 384
DB 639 CACACTTCATTAC 626

RESULT 2 987 bp DNA linear GSS 26-JUN-1999
CNS014PO
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN12P22 of Drosophila library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL104456.1 GI:5616067

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 987)

REFERENCE
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (23-JUN-1999) Genoscope - Centre National de Sequencage ;
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
<http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billand at CEPH (Centre
d'Etude de Polymorphism Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

COMMENT

FEATURES
source
1..987
/location/Qualifiers
/organism="Drosophila melanogaster"

/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN12P22"
/clone_1ib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : SP6"

ORIGIN

Query Match 9.0%; Score 46; DB 14; Length 987;
Best Local Similarity 33.5%; Pred. No. 0.59;
Matches 78; Conservative 51; Mismatches 104; Indels 0; Gaps 0;

QY 136 ACTATTAATGCTTATGCAACAAAGAAATGCAATATCATCTTACTGTCATATATAT 195
DB 723 AAAAAAAMAAAMAAAMTWTATTAATTAATTAATTAATTAATTAATTAATTAAT 782
QY 196 TTCAATTAATGCGGTTTATTAATCACTCAATTAATTAATTAATTAATTAATTA 255
DB 783 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 842
QY 256 GTGACTGTTTAAAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 315
DB 843 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 902
QY 316 GAGGATTCAGCAACAGTGTCTTACAACTGAAAGCAATTAATTAATTAATTA 368
DB 903 RAAMWMMWMTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 955

RESULT 3 1101 bp DNA linear GSS 04-JUN-1999
CNS00E0E/c
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR28A22 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL067626
AL067626.1 GI:4948726

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazuo Oosawa and
Aaron Mammeter in Pictet de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR28A22"
/clone_1ib="RPCI-98"
/note="end : TET3"

ORIGIN

	Query Match	Score	DB	Length	Best Local Similarity	Pred. No.	Matches	Conservative	Mismatches	Indels	Gaps
OY	130	CCCACCACTPAAGTCCTATGCAACAAAGAAATGCAATACATCTCTTAGTCTCATT	189								
Db	1100	CCCAACSCAPARARAHGNSADDAADAAADKAAALACTTTTTTTTTTTTTTTTTTTTTT	1041								
OY	190	ATTATTTTCATTAGTAGCGGTTTTTTTACTACATCTCAATTAAGTAGACGAATGANT	249								
Db	1040	TTTTTTTTTTTATTSAAATCCCMATCAAAAACGCGCBCCASCAATTTACAAATTAAGCAW	981								
OY	250	GAGTTAGTACCTGTTATPAAGAGAGTAAATPAAGATCTATCATCATTTGAGGCAATPA	309								
Db	980	TTCTWAKKTCMAMAMAMAMAMAMAAATAGTAGGAAATCTGAGGGAATAAGCAAGAAA	921								
OY	310	GAGAGGAGGAG	318								
Db	920	GCRACGGG	912								

RESULT 4			
LOCUS	AZ541558/c		
DEFINITION	AZ541558	910 bp	DNA linear GSS 14-NOV-2000
ACCESSION	ENTEB157P	Entamoeba histolytica	Sheared DNA Entamoeba histolytica
VERSION	AZ541558	genomic, genomic survey sequence.	
KEYWORDS	AZ541558.1	GI:11149198	
SOURCE	GSS.		
ORGANISM	Entamoeba histolytica		
REFERENCE	Entamoeba histolytica		
AUTHORS	Bukaryota; Entamoebidae; Entamoeba.		
TITLE	1 (bases 1 to 910)		
JOURNAL	Loftus,B., Van Aken,S. and Fraser,C.		
COMMENT	Determination of clone end sequences from Entamoeba histolytica		
	HMI:INSS sheared DNA library		
	Unpublished (2000)		
	Contact: Brendan J Loftus		

FEATURES
SOURCE

/organism="Entamoeba histolytica"
 /mol_type="genomic DNA"
 /strain="HM1:IMSS"
 /db_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /note="Vector: pHO3; Site_1: Bst II. Constructed at The
 Institute for Genomic Research (TIGR), Rockville, MD.
 Genomic DNA isolated from broth cultures of *E. histolytica*
 using a method described by Clark and Diamond (Clark,
 C.G., and Diamond, I.S. (1993) *Entamoeba histolytica*: a
 method for isolate identification. *Exp. Parasitol.*
 77:450.). The DNA was mechanically sheared to give a
 tight size distribution (~2 kb). The v + i method used for
 the library construction is described in detail in Smith,
 H.O. and Venter, J.C. (Making small insert libraries for
 whole genome shotgun sequencing projects. In *Genome
 Sequencing: A Practical Approach*, eds. M. Vaudin and B.
 Jarell, Oxford University Press, 1999).

ORIGIN

Query Match	8.7%	Score 44.4	DB 11	Length 910
Best Local Similarity	46.7%	Pred. No. 1.5		
Matches 141	Conservative 0	Mismatches 161	Indels 0	Gaps 0
QY	37	AGAAATGATTCACCAAAATAGATGACAAAGACAGATTTGTGTTTATCTGTGGG	96	
DB	310	AGAAATAAAATGAAATATAGATPAAATATGAAATGAAATTTATTTTATTTATTTCAATGAT	251	
QY	97	TAAATAGCTTTCTCCAGTTGTATPAAAGACCCCTCCACAGTATPAAAGTCTTATGCACA	156	
DB	250	TAAATTCATTAAGATPAAATGAAATAGATATATAGAAATPAAAGAAATPAAACAAATPAAAGAAAT	191	
QY	157	AGAAGATGTCATATACATTCCTCTAGTCTCATTTATTTTCATTTAGATAGCCGGTTTTT	216	
DB	190	GATTAATATACATCAACATATTTATGTATTTATTTCTTATTTCTTATTTATTCATTTCTTATTTTC	131	
QY	217	TACTACAACTCAATTAAGATGACAGATGATGGGTAGCTGTTTATPAAAGACAG	276	
DB	130	TATTTAAACATATATTAACAAACATATATAGAAAGATPAAATPAAAGATPAAACATCATPAA	71	
QY	277	TAAATPAAAGATTCATCATCTATTGAGGCATPAAAGGAGGAGAGATTCAGCAACAGTGT	336	
DB	70	AATPAAAGATPAAAGATPAAATATATCAACAAACATATPAAACAAAGAAATPAAACCAATACTTT	11	
QY	337	GC 338		
DB	10	CC 9		

RESULT 5	BH155022	894 bp	DNA	linear	GSS 24-SEP-2001
LOCUS	BH155022				
DEFINITION	BH155022	ENT51357F	Entamoeba histolytica	Sheared DNA	Entamoeba histolytica
ACCESSION	BH155022				
VERSION	BH155022.1	GI:15726739			
KEYWORDS	GSS.				
SOURCE	Entamoeba histolytica				
ORGANISM	Entamoeba histolytica				
REFERENCE	1 (bases 1 to 894)				
AUTHORS	Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.				
TITLE	Determination of clone and sequences from Entamoeba histolytica				
JOURNAL	HMJ:IMSS sheared DNA library (2001)				
COMMENT	Unpublished (2001)				
CONTACT	Contact: Brendan J Loftus				

FEATURES

SOURCE

```

/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM: IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOsi; Site 1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
right size distribution (~2 kb). The v + i method used for

```


Db 978 RGGRRKDKKORRKKDDKGGKKKKAAKAAKAAATKMWDDWMDKMWKDKGAKDRKADD 919
Qy 238 AACAGATGAATGGGTTAGTGACTGTTTAAAGAGATTAATGAATATCATCAT 297
Db 918 DGAAGDKDDGKADDDDTGKTDDDKKDDMKAGTGDATWAAATDMMWGMWA 859
Qy 238 TTGAGGCATTAAGGAGGAGAGATTCAACAAACAGTGTGCTTACAAGTGAACAAAGT 357
Db 858 DADWMTWDAADDMWADDRWDAMWAMKMDAMAGATRADRDGDAKGGARKKRRDK 799
Qy 358 TTAACCTAAAGTGA 370
Db 798 RADDKRDAADRD 786

RESULT 8
CL108838/c CL108838 1158 bp DNA linear GSS 05-JAN-2004
LOCUS ISB1-5012, Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-5012,
DEFINITION genomic survey sequence.
ACCESSION CL108838
VERSION CL108838.1 GI:40602473
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 1158)
Kremliuk, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert length: 75000 Std Error: 0.00
Seq primer: Sp6 ATTTAGTGACACTATAG
Class: BAC ends
High quality sequence start: 36
High quality sequence stop: 766.
Location/Qualifiers
1..1158
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-5012"
/clone_1fb="ISB1"
/note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC
library Segment 1"

ORIGIN
Query Match 8.6%; Score 43.8; DB 12; Length 1158;
Best Local Similarity 52.5%; Pred. No. 2.3;
Matches 96; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Db 115 TTGTATTAAGACCTCCACAGATTAAGTCTATGACAAAGAAAAATGCAATACAT 174
Qy 781 TGGATTAAGGCACTTCACAGAAATAGTAGCTCAAAAAAGAAAAAGAAAAA 722
Db 175 TCTCTAGTCATTAATTTTCAATTAATAGCCGGTTTTCATGACAACTCAATAG 234
Qy 721 GCTGTTGGGTAATTTGTAATGTTTATAGAGCCGTGAAGACCTAAGTAAGAAA 662
Db 235 ATGAACAGATGATGCGTTAGTGACTGTTTAAAGAGAGATTAAGATATCAT 294
Qy 661 GTTAAATATCTAAATGTTATGCACTATATCCCAATAGTAAGTAATTAATTA 602
Db 295 CAT 297
Qy 601 TAT 599

RESULT 9
CG769354 927 bp DNA linear GSS 29-OCT-2003
LOCUS TcB41.1_H04_Sp6 Tribolium BAC library Tribolium castaneum genomic,
DEFINITION genomic survey sequence.
ACCESSION CG769354
VERSION CG769354.1 GI:38022533
KEYWORDS GSS.
SOURCE Tribolium castaneum (red flour beetle)
ORGANISM Tribolium castaneum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
Tenebrionidae; Tribolium.
1 (bases 1 to 927)
Savard, J. and Tautz, D.
Tribolium castaneum BAC-ends sequencing project
Unpublished (2003)
Contact: Savard, J.
Abteilung fur Evolutionsgenetik, AG Tautz
Institut fur Genetik, Universitat zu Koln
Weyertal 121, 50931 Koln, Germany
Tel: 49 221 470 6911
Fax: 49 221 470 5975
Email: savard@uni-koeln.de
Class: BAC ends.

FEATURES
Location/Qualifiers
1..927
/organism="Tribolium castaneum"
/mol_type="genomic DNA"
/strain="GA-2"
/db_xref="taxon:7070"
/clone_1fb="Tribolium BAC library"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
library constructed by Exelixis Inc."

ORIGIN
Query Match 8.5%; Score 43.6; DB 12; Length 927;
Best Local Similarity 53.8%; Pred. No. 2.5;
Matches 112; Conservative 0; Mismatches 94; Indels 2; Gaps 1;

Qy 138 TTTAAAGTCTATGCAACAAGAAATGCAATATCTCTAGTCTCATTTATTTT 197
Db 26 TATTAAGTACGTTGTAATTTAAATGTCATTAATTTGTTACAAATGCAACTA 85
Qy 198 CATTAAGATCCCGGTTTCTACTACACTCAATTAAGATGAACAGATGAATGGTTAGT 257
Db 86 AATTCATTTTACACTGTACAGAAATATCTTTTATTTTAACAAAAAGAAACCGATATT 145
Qy 258 GACTGTTTATTAAGAAGATATTA--GATCTATCATCATTTGAGCAATTAAGGAGG 315
Db 146 AAACGCATTAACAGATTAATTAACCGATACCATTTATTAAGTCAAGCTGTAGG 205
Qy 316 GAGAGATTCAGCAACAGTGTCTTACA 343
Db 206 AAGAGATTACAGAAACCGCTGTTTCA 233

RESULT 10
AV008603 443 bp mRNA linear EST 25-AUG-1999
LOCUS AV008603 Mus musculus 18-day embryo C57BL/6J Mus musculus cDNA
DEFINITION clone 1110013A05, mRNA sequence.
ACCESSION AV008603
VERSION AV008603.1 GI:4785590
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Mus.
1 (bases 1 to 443)

AUTHORS Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K., Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara, A., Hayata, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuna, H., Oda, H., Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tateo, M., Tomaru, Y., Tomimaga, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

TITLE RIKEN Mouse ESTs

JOURNAL Unpublished (1999)

COMMENT Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadaei, Tsukuba, Ibaraki 305-0074, Japan
Tel.: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-resortc.riken.go.jp
Thermolabile and thermostable activation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES
source
1..443
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="1110013A05"
/sex="mixed"
/dev_stage="18-day embryo"
/clone_11b="Mus musculus 18-day embryo C57BL/6J"

ORIGIN
Query Match 8.4%; Score 43.2; DB 1; Length 443;
Best Local Similarity 55.3%; Pred. No. 2.8;
Matches 84; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

154 ACAAGAAATGTCATTCATCTCTAGTCATATTTATTTATAGATGAGCGGTT 213
3 ACTTAATTTCTCTCAACATCCCTTAATTTTCATTTAGATGCGTGAAGAA 62
214 TTTTACTCAACTCAATTAAGATGAACAGATGATGGTGTAGTCTTTATTAAGA 273
63 ATGAACCTTCAGAGAACAGAACATGAATAAGTGGTTTATGACGTGTTTGAGATTA 122
274 GAGTATATAAGTATCATCATCATTTGAGGCA 305
123 AAGATAGGTGCTCATAGAGTTTATGATTA 154

RESULT 11
LOCUS DX034017 784 bp DNA linear GSS 10-JAN-2006
DEFINITION KBRB031L03F KBRB, Brassica rapa BamHI BAC library Brassica rapa subsp. pekinensis genomic clone KBRB031L03, genomic survey sequence.
ACCESSION DX034017
VERSION DX034017
KEYWORDS DX034017.1 GI:84728314
ORGANISM GSS.
Brassica rapa subsp. pekinensis
Brassica rapa subsp. pekinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eustosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 784)
Yang, T.J., Kwon, S.J., Kim, J.A., Kim, J.S., Lim, K.B., Jin, M., Park, J.Y., Lim, M.H., Kim, H.I., Choi, B.S., Seol, Y.J., Park, D.S., Hahn, J.H. and Park, B.S.
End sequence of Brassica rapa BamHI (KBRB) BAC clone unpublished (2005)

COMMENT Contact: Beom-Seok Park
Brassica Genomics Team
National Institute of Agricultural Biotechnology
225 Seodun-Dong, Suwon, 441-707, Korea
Tel.: +82-31-299-1670
Fax: +82-31-299-1672
Email: pbeom@da.wo.kr
BAC end sequence of Brassica rapa subsp. pekinensis BamHI BAC clone KBRB031L03
Seq primer: 77
Class: BAC ends.

FEATURES
source
1..784
/organism="Brassica rapa subsp. pekinensis"
/mol_type="genomic DNA"
/cultivar="Chilifu"
/sub_species="pekinensis"
/db_xref="taxon:51351"
/clone="KBRB031L03"
/lab_host="E. coli DH10B"
/clone_11b="KBRB, Brassica rapa BamHI BAC library"
/note="Vector: pCUGIBAC1; Site 1: BamHI; Brassica rapa subsp. pekinensis var. Chilifu BAC library (KBRB BAC) is provided by Yong-Pyo Lim (CNU)."

ORIGIN
Query Match 8.4%; Score 43.2; DB 14; Length 784;
Best Local Similarity 50.0%; Pred. No. 3.1;
Matches 108; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

58 AGAATGACAAAGCAGAGATTGTGTTTATCTGTGGTAAATACGTTTCTCCAGTTG 117
368 AAAATCGCGCCAGACAGATTAGGTTTATTTTATTAATTAAAGACATAGTTACGGAA 427
118 TATTAAGACCTCCACAGATTAAGTCTATGCAACAAAGAAATGTCAATCATCTT 177
428 CAAATATATATTAATTAATCAAAATTAATAAGCTATTAATATGTCACATTAATTTT 487
178 CTTAGTCTCATTTATTTTATTTATTTATTTATTTTATTTTATTTTATTTTATTTT 237
488 TCACATGCTATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 547
238 AACGAATGAATGGGTAGTACTGTTTATTAAGA 273
548 AACATATTAATAATGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 583

RESULT 12
LOCUS DX077691 886 bp DNA linear GSS 10-JAN-2006
DEFINITION KBRB089F21F KBRB, Brassica rapa BamHI BAC library Brassica rapa subsp. pekinensis genomic clone KBRB089F21, genomic survey sequence.
ACCESSION DX077691
VERSION DX077691
KEYWORDS DX077691.1 GI:84771987
ORGANISM GSS.
Brassica rapa subsp. pekinensis
Brassica rapa subsp. pekinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eustosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 886)
Yang, T.J., Kwon, S.J., Kim, J.A., Kim, J.S., Lim, K.B., Jin, M., Park, J.Y., Lim, M.H., Kim, H.I., Choi, B.S., Seol, Y.J., Park, D.S., Hahn, J.H. and Park, B.S.
End sequence of Brassica rapa BamHI (KBRB) BAC clone unpublished (2005)
Contact: Beom-Seok Park
Brassica Genomics Team
National Institute of Agricultural Biotechnology
225 Seodun-Dong, Suwon, 441-707, Korea
Tel.: +82-31-299-1670
Fax: +82-31-299-1672

TITLE Unpublished (2005)

JOURNAL Contact: Beom-Seok Park
Brassica Genomics Team
National Institute of Agricultural Biotechnology
225 Seodun-Dong, Suwon, 441-707, Korea
Tel.: +82-31-299-1670
Fax: +82-31-299-1672

DB 234 ATAAATGACCTTTGTTTGTGTTGATATATACGTTTGAAGTTAATTTTATTA 293
QY 229 AATAAGATGACAGAAATGAGGTTAGTACCTGTTTAAAGAGACTTAATAAGATAC 288
DB 294 TCAACACAAATTCGCAAGAAATAGATGTTGGGATCCATATTTTGGCAAAATATATATAC 353
QY 289 TATCATCTTTGAGGCAATTAAGGAGAGATTCACCAACAGCTGTGCTTACAGTGG 348
DB 354 TTGCTTCAATTTAAAGTTAAAGACACGTTAAATATAGCAAAAATGCTATGCAATTTGG 413
QY 349 AAACAGCTTAATAAGTGAACCCCTCTCTTGCAAGATCAATGCGCAGTTG 403
DB 414 CCCATTAGTTAAACAAATATAGTTGCTCATTAACAAATTAAGCCCAATTGG 468

RESULT 15
BH137539/c 879 bp DNA 1linear GSS 07-AUG-2001

LOCUS ENTMA831R Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, genomic survey sequence.

ACCESSION BH137539
KEYWORDS BH137539.1 GI:15096600

SOURCE GSS.
Entamoeba histolytica

ORGANISM Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.

REFERENCE 1 (bases 1 to 879)

AUTHORS Loftus, B., Wang, Z., Van Aken, S. and Fraser, C.
TITLE Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library (2001)

JOURNAL Unpublished (2001)

COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@tigr.org

Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library

Seq primer: M13-Reverse

Class: shotgun
High quality sequence start: 32
High quality sequence stop: 858.
Location/Qualifiers

1.879
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOSt1; Site 1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."

FEATURES
source

ORIGIN

Query Match 8.4%; Score 43; DB 11; Length 879;

Best Local Similarity 49.8%; Pred. No. 3.5;

Matches 109; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 148 TATGCAACAAAGAAATGTCATATCTTCTTATGCTCATATATATTTTCATTAGATAG 207

DB 303 TACAAAAGAGTTAATTCACCTTGATTAATGTTTTATGAAATGATATTCAG 244
QY 208 CCGGTTTTTACTRCAACTCAATTAAGATGAACGAATGAATGGTTAGTACTGTTAT 267
DB 243 AAGTGAAATACCAACAAAGTGAATGAATGAATGAATGAATGATGTTTGTGAATGTTCA 184
QY 268 AAGAGAGTAAATTAAGATATCTATCATCTTGGAGCAATTAAGAGAGAGATTCAGC 327
DB 183 CATTAACAAATTAATATATATCATCATCATCATCATCATCATCATCATCATCATCATCAT 124
QY 328 AACAGTGTCTTACAAAGTGAAGAAACAAAGTTAACTTAA 366
DB 123 ATTGTGATGTGAAGAAATTAAGAAATTAAGAAATTAAGAA 85

Search completed: May 28, 2006, 14:46:43
Job time : 2628.13 secs

Db 350 NNNNNNTYKKBHBMWVHSAWKHNAHYSRKKTBRYKKTWNNNGTTMWR 409
QY 312 GAGGAGAGATTGAGCAACAGTGTCTTACAGTGAAGAAACAAGTTAACTAAAGTGAC 371
Db 410 MAAWYWKMDMBWGTNNNNNGRTYYGWTYKXKMYKWKYANNCKRABMDHKTCTHNN 469
QY 372 CCCCCCTCTTGCAGATCAATGACAGATTGAGCTTACGACGACACATCATCTGTAA 431
Db 470 TTMWOKTYMNNCKYKSMKTNGSHBBAALVYTWYMMWRRYAHANNNNNDYMMRACCTWYK 529
QY 432 ATGCTTTCCTGATAGACCTGTTCAATTAATCTCTTTGCAAGC 475
Db 530 YBVCCKMWNMYAAWYTKSSWNTSYRYMKTNNSWRMRSDTRSM 573

RESULT 2

US-09-270-767-15058/C
Sequence 15058, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 15058
LENGTH: 1257
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-15058

Query Match 8.2%; Score 42.2; DB 3; Length 1257;
Best Local Similarity 53.3%; Pred. No. 0.033;
Matches 89; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 73 GAAATTTGCTTTTATCTGTGGTAATAAGTTTTCTCAAGTTGATTAAGACCTCCC 132
Db 364 GAGTTTATGATTTAATTTAATTTAATTTCACTCAACACAGACACGACCTACTC 305
QY 133 ACCAGTATAAGTCTATGCAACAAGAAAGTCAATCTCTTCTGTTCTCATTTT 192
Db 304 ACAGCAAAAACGTACAAGAGAGAAAGAAATTAAGAGGTATTTCTTCAATA 245
QY 193 ATTTTATGATGATGCGGCTTTTATCTCAACTCAATTAAGATGA 239
Db 244 TGTATTGCAATAAAGGTGTGCCATTCATATCAATATAAAGTA 198

RESULT 3

US-09-270-767-10136/C
Sequence 10136, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10136
LENGTH: 733
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-10136

Query Match 7.8%; Score 40; DB 3; Length 733;
Best Local Similarity 50.5%; Pred. No. 0.12;
Matches 97; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
QY 48 TCACCAATTGAAATGACAAAGAGATTGTTTATCTGTGGTAATAAGCTTT 107

Db 306 TAAATATATGATGAGTTTATTTAGTTTATTTATTTAAGTTATTTCACT 247
QY 108 TCTCCAGTTGTATTAAGACCTCCCAAGTATTAAGTCTATGCAAAAGAAATGTC 167
Db 246 GCAACACGACGACGACCTACTACGACAAAGAAACGTACAAAGAGAAAGAAATA 187
QY 168 AATACATCTCTTACTGCTCATATATTTTTCATATAGATAGCGGTTTTTACTCACTC 227
Db 186 AAAAAAGGTATTTCTCTTCAATATGTTTATATGCAATAAAGGTGTGCCATTCATATC 127
QY 228 AATATAGATGA 239
Db 126 AATATATAAGTA 115

RESULT 4

US-09-270-767-10504/C
Sequence 10504, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10504
LENGTH: 4860
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-10504

Query Match 7.7%; Score 39.2; DB 3; Length 4860;
Best Local Similarity 46.6%; Pred. No. 0.39;
Matches 125; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 97 TAAATAGTTTTCTCCAGTTGTATTAAGACCTCCACAGTATTAAGTCTATGCAACA 156
Db 1390 TAAATAGTTATTTTGTATTAATATCAACGATTAATATGGAAGCATTAATTT 1331
QY 157 AAGAAATGCAATACATCTCTAGTCTCATTTATTTTTCATATGATAGCGGTTTTT 216
Db 1330 AATTAATTTATTTTATTAATATGCGCATATGCTTAATAATTTTTTGGCTTACA 1271
QY 217 TACTACACTCAATATAGATGAACAGATGAATGGTTAGTACTGTTATTAAGAGAG 276
Db 1270 TTTTGTAGCAATTAATAATTTTAATTTACATTTGATAGCTTTGATTTTAAGATGC 1211
QY 277 TAATAAGATGATCATCATTTTGAAGCAATTAAGGAGAGAGATTGAGCAAGTGT 336
Db 1210 CAATAATGTTTTTATCTTAATATGAGTAAGAAAGATTAATAAACCAATTAAGTTC 1151
QY 337 GCTTACAGTGAACCAAGTTAACTA 364
Db 1150 ATTTTATGTTAAATTAAGTACGTGA 1123

RESULT 5

US-09-662-254B-27
Sequence 27, Application US/09662254B
Patent No. 6933145
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Li, Yi
TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous
FILE REFERENCE: UF-221C1X1
CURRENT APPLICATION NUMBER: US/09/662,254B
CURRENT FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 09/086,651

;; PRIOR FILING DATE: 1998-05-29
;; PRIOR APPLICATION NUMBER: 60/224,479
;; PRIOR FILING DATE: 2000-08-10
;; NUMBER OF SEQ ID NOS: 80
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 27
;; LENGTH: 32392
;; TYPE: DNA
;; ORGANISM: Ambacta moorei entomopoxvirus
US-09-662-254B-27

Query Match 7.5%; Score 38.4; DB 3; Length 32392;
Best Local Similarity 46.3%; Pred. No. 1.3;
Matches 126; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 19 ATGCTATTATGAGTGAATGATTTCTCACCAATTGAAATGACACAGAGATT 78
DB 23294 ATGTTATGATGATGATTTCTCAATAAGCGAAATATATATAGAAAAGTGACT 23353
QY 79 TGTGTTTATCTGTTGGTAATACGTTTCTCCAGTTATTAAGACCCCTCCACAGT 138
DB 23354 GGTGTCCTCTATGCTTATATATATATATATATATATATATATATATATAT 23413
QY 139 ATAAAGTCTATGCAACAAAGAAATGCAATACATTTCTTATGCTCATTTATTTTC 198
DB 23414 AAAAGGAACAAATGATGAAATATATATATATATATATATATATATATATAT 23473
QY 199 ATTAATAGCCGGTTTTTTTACTCACTCAATTAAGATACAGATGAAATGGTTAGT 258
DB 23474 ATTAATACATAAGATGAAATTTAAACATATATATATATATATATATATATAT 23533
QY 259 ACTGTTTATTAAGAGATATATAGATATCTA 290
DB 23534 ATGCTTTAAATATATATTAACAAATATATA 23565

RESULT 6
US-09-662-254B-23/C
; Sequence 23, Application US/09662254B
; Patent No. 6933145
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Li, Yi
; APPLICANT: Bawden, Alison Louise
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous
; FILE REFERENCE: UF-221C1XCL
; CURRENT APPLICATION NUMBER: US/09/662,254B
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 09/086,651
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/224,479
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Ambacta moorei entomopoxvirus
US-09-662-254B-23

Query Match 7.5%; Score 38.4; DB 3; Length 50000;
Best Local Similarity 46.3%; Pred. No. 1.5; Indels 0; Gaps 0;
Matches 126; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
QY 19 ATGCTATTATGAGTGAATGATTTCTCACCAATTGAAATGACACAGAGATT 78
DB 9099 ATGTTATGATGATGATTTCTCAATAAGCGAAATATATATAGAAAAGTGACT 9040
QY 79 TGTGTTTATCTGTTGGTAATACGTTTCTCCAGTTATTAAGACCCCTCCACAGT 138
DB 9039 GGTGTCCTCTATGCTTATATATATATATATATATATATATATATATATATAT 8980

QY 139 ATAAAGTCTATGCAACAAAGAAATGCAATACATTTCTTATGCTCATTTATTTTC 198
DB 8979 AAAAGGAACAAATGATGAAATATATATATATATATATATATATATATATAT 8920
QY 199 ATTAATAGCCGGTTTTTTTACTCACTCAATTAAGATACAGATGAAATGGTTAGT 258
DB 8919 ATTAATACATAAGATGAAATTTAAACATATATATATATATATATATATATAT 8860
QY 259 ACTGTTTATTAAGAGATATATAGATATCTA 290
DB 8859 ATGCTTTAAATATATATTAACAAATATATA 8828

RESULT 7
US-09-782-378A-17/C
; Sequence 17, Application US/09782378A
; Patent No. 6916635
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Nadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatlenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; PRIOR FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 56737
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-17

Query Match 7.5%; Score 38.2; DB 3; Length 56737;
Best Local Similarity 52.9%; Pred. No. 1.8;
Matches 82; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 156 AAAAGAAATGCAATACATTTCTTATGCTCATTTATTTTCACTTATGACCGGTTTT 215
DB 18242 AAATAAATATGTCMAAGCAATTTTCTACCTCAGATGATCTTATATACATGAAATGCTTT 18183
QY 216 TTACTACAACTCAATTAAGATGACAGAAATGAGTTAGTCACTGTTATTAAGAGA 275
DB 18182 TTAAGAAATATATTAAGTTAACTCCCTATTTTGTCTGTTTGTCTTATCTAATAA 18123
QY 276 GTAATTAAGATATCTATCATTTGAGGCAATAG 310
DB 18122 ACATTCGACAAATCCCAAGATTAATGTCATAGC 18088

RESULT 8
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEKI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-1071160
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA

ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 7.4%; Score 38; DB 3; Length 640681;
Best Local Similarity 46.6%; Pred. No. 5;
Matches 122; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

```
OY 24 TATTATGAGTAGATAGATTCACCAATAGATGACAAAGCAGAGATTGTGT 83
DB 189923 TATTATGAGTAGATAGATTCACCAATAGATGACAAAGCAGAGATTGTGT 189982
OY 84 TTTATCTGTGGTAATATACGTTTCTCCAGTTGTATTAAGACCCCTCCACAGTATATA 143
DB 189983 TATAGCAAAATTAGAAATTTCTCCGGAATATATTAAAAATAAGAAAAAAATTA 190042
OY 144 GTCCATGCAACAAAGAAATGCAATATCATCTCTAGTCTCATTTATTTTCAATTAG 203
DB 190043 TTTTATTTTAAACAAATAATTAATAATTTTATTTTATTTTATTTTATTTTAA 190102
OY 204 ATACCCGTTTCTTACTACACTCAATTAAGATGACAGATGATGGTTAGTACTGT 263
DB 190103 TTCTCTTTCTAGAAAAACAATATGATTTTATTTTAAAGTTGATGTTTAAAAAAT 190162
OY 264 TTATAAGAGAGATTAATAAGA 285
DB 190163 AAAAAATGAAACAAAAACAATA 190184
```

RESULT 9

US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367

GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOMLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29, 768
REFERENCE/DOCKET NUMBER: 30472/114 IMM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-9300
TELEFAX: (703) 883-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
CLONE: PTZ9PC-F18
US-08-232-463-14

Query Match 7.4%; Score 37.8; DB 2; Length 7218;
Best Local Similarity 13.5%; Pred. No. 1.1;
Matches 30; Conservative 103; Mismatches 90; Indels 0; Gaps 0;

```
OY 148 TATGCAACAAAGAAATGCAATATCATCTCTTAGTCTCATTTATTTTCAATAGATAG 207
DB 1494 TAGGATCACTGTAATATCATATCATATGATTAAGAGATTAAGAAATTTGGTAC 1435
OY 208 CCGGTTTTTACTACAACTCAATTAAGATGAACAAATGAATGGTTAGTACTGTTAT 267
DB 1434 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1375
OY 268 AAAGAGAGTAATAAGATCATCATCATCTTTAGGGAATTAAGGAGAGAGATTAAC 327
DB 1374 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1315
OY 328 AAACAGTGTCTTACAAAGTGAAACAAAGTTAACTAAAGTGA 370
DB 1314 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1272
```

RESULT 10

US-09-949-016-125489/c
Sequence 125489, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 125489
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-125489

Query Match 7.3%; Score 37.6; DB 3; Length 601;
Best Local Similarity 49.0%; Pred. No. 0.54;
Matches 100; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

```
OY 148 TATGCAACAAAGAAATGCAATATCATCTCTTAGTCTCATTTATTTTCAATAGATAG 207
DB 365 TATTAATATATAGTATATATCAACAGATGATATATTAATCATGATTAATTAACATAT 306
OY 208 CCGGTTTTTACTACAACTCAATTAAGATGAACAAATGAATGGTTAGTACTGTTAT 267
DB 305 ACATRTATATAGTATGATTAATCTGAAGAGATGCAAAATTTCTGCTTACTATCTTA 246
OY 268 AAAGAGAGTAATAAGATCATCATCATCTTTAGGGAATTAAGGAGAGAGATTAAC 327
DB 245 TTAATGACATCTTATTAAGTCTCTCAAAATTTAAGATTTAAATATTTCAAGCTCAGT 186
OY 328 AAACAGTGTCTTACAAAGTGAA 351
DB 185 CAACATTAATGCTTAAAGTAAAAA 162
```

RESULT 11

US-09-949-016-15302/c

; Sequence 15302, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 15302
; LENGTH: 44353
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(44353)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15302

Query Match 7.3%; Score 37.6; DB 3; Length 44353;
Best Local Similarity 49.0%; Pred. No. 2.5;
Matches 100; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

Qy 148 TATGACAAAGAAATGTCATATCATCTCTTACGTCATTAATTTCTTACATGATG 207
Db 25736 TATTAATAATATGATATATATCAACAGATATATTAACATGATATATGTTAACTACATAT 25677
Qy 208 CCGGTTTTTACTACAACTCAATAGATGAACAGATGATGGTGAAGCTGTTAT 267
Db 25676 ACATGATATATAGTATGATTAATCTAGAAAGATGAAAGATTTCTCGCTACTATCTGA 25617
Qy 268 AAGAGAGATTAATAAGATCATCTATCTTTGAGCATTAAGGAGAGAGATTCACG 327
Db 25616 TTAATGCAATCTTAACTCTCAAAATATTAAGAAATTAATTCAAAGCGTCACT 25557
Qy 328 AAACAGTGTCTTACAGTGA 351
Db 25556 CAACATATGCTTAAAGTAAAA 25553

RESULT 12
US-09-531-120-211
; Sequence 211, Application US/09531120
; Patent No. 6972197
; GENERAL INFORMATION:
; APPLICANT: PREUSS, DAPHNE
; APPLICANT: COPENHAVER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/09/531,120
; CURRENT FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
; LENGTH: 1082144
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-531-120-211

Query Match 7.3%; Score 37.6; DB 4; Length 1082144;
Best Local Similarity 50.0%; Pred. No. 7.8;
Matches 94; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy 18 CATGCTTATATGAGAGATAGATTCACCAAAATTAGATGACAAAGCAGAT 77
Db 431379 CACACATATTAATAAAATCATTTAAAAATATTTTCACACAGATTAAGAAATCATTTAA 431438
Qy 78 TTGTGTTTATCTGTGGTAATACGTTTCTCCAGTTGTAATAAGACCTCCACCG 137
Db 431439 TTTTATTTTATTTGTAAGATTAACATCTTTATTTATTTCAATTAATGATTCACCA 431498
Qy 138 TATTAAGTCTTATGACAAAGAAATGTCATATCTCTTACGTCATTAATTTT 197
Db 431499 TAAAAATATATCTGCGAATATTAATCGCAAAATTAATAAAATTAATATTCCTTA 431558
Qy 198 CATTAGAT 205
Db 431559 ATTATAT 431566

RESULT 13
US-08-998-416-187/c
; Sequence 187, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Redlschning, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPPII
; NUMBER OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
; COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
LENGTH: 663 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1074UP
US-08-998-416-187

Query Match 7.3%; Score 37.4; DB 3; Length 663;
Best Local Similarity 58.6%; Pred. No. 0.63;
Matches 65; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200
QY	PAGTCGTCAATTAATTTTCATAGATAGCGCGGTTTTTCTCAACATCGAAATGAATGAA	239																			
Db	574	TGATTTATTAATTAACCTTTATTTAGTTAAACCATTTATTTATGATCATATAATATATATAA	515																		
QY	240	CAGATGAAATGGGTTAGTACTCTTTATAAAGAAAGTATTAATGAATACATA	290																		
Db	514	GGATATACCTTAATGATATATATAGTTATTTATAGAACCAATGAAGATACATA	464																		

RESULT 14
US-08-998-416-779
; Sequence 779, Application US/08998416
; Date of Issue 03/03/2004

```

1  GENERATED INFORMATION:
2  APPLICANT:  PHILIPSEN, Peter
3  APPLICANT:  Pohlmann, Rainer
4  APPLICANT:  Steinert, Sabine
5  APPLICANT:  Mont, Christine
6  APPLICANT:  Wendland, Jürgen
7  APPLICANT:  Knechtle, Philipp
8  APPLICANT:  Reischung, Corinne
9  TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSEYPTII
10 TITLE OF INVENTION:  AND USES THEREOF
11 NUMBER OF SEQUENCES:  1152
12 CORRESPONDENCE ADDRESS:
13 ADDRESSEE:  No. 6239264artis Corporation
14 STREET:  3054 Cornwallis Road
15 CITY:  Research Triangle Park
16 STATE:  No. 6239264eth Carolina
17 COUNTRY:  USA
18 ZIP:  27709
19
20 COMPUTER READABLE FORM:
21 MEDIUM TYPE:  Floppy disk
22 COMPUTER:  IBM PC compatible
23 OPERATING SYSTEM:  PC-DOS/MS-DOS
24 SOFTWARE:  Patent in Release #1.0, Version #1.30
25
26 CURRENT APPLICATION DATA:
27 APPLICATION NUMBER:  US/08/998,416
28 FILING DATE:  24-DEC-1997
29 CLASSIFICATION:  415
30
31 PRIOR APPLICATION DATA:
32 APPLICATION NUMBER:  CH 0016/97
33 FILING DATE:  31-DEC-1996
34 ATTORNEY/AGENT INFORMATION:
35 NAME:  Meigs, J. Timothy
36 REGISTRATION NUMBER:  38,241
37 REFERENCE/DOCKET NUMBER:  Pf/5-30306/A/CGC1976
38
39 TELECOMMUNICATION INFORMATION:
40 TELEPHONE:  919-541-8587
41 TELEFAX:  919-541-8689
42
43 INFORMATION FOR SEQ ID NO:  779:
44 SEQUENCE CHARACTERISTICS:
45 LENGTH:  696 base pairs
46 TYPE:  nucleic acid
47 STRANDEDNESS:  single
48 TOPOLOGY:  linear
49
50 MOLECULE TYPE:  DNA (genomic)
51
52 ORIGINAL SOURCE:
53 ORGANISM:  P4G1501RP
54
55 US-08-998-416-779

```

Query Match	7.3%	Score 37.4	DB 3	length 696
Similarity				
Best Local	58.6%	Pred. No. 0	64	
Matches	65	Conservative	0	Mismatches 46
			Indels 0	Gaps 0

[illegible]

RESULT 15
US-08-998-416-1138/c
; Sequence 1138, Application US/08998416
; Patent No. 6239264

? APPLICANT: Philippsen, Peter
 ? APPLICANT: Pohlmann, Rainer
 ? APPLICANT: Steiner, Sabine
 ? APPLICANT: Mohr, Christine
 ? APPLICANT: Wendland, Jürgen
 ? APPLICANT: Knechtle, Philipp
 ? APPLICANT: Reischung, Corinne
 ? TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSEYPTII
 ? TITLE OF INVENTION: AND USES THEREOF
 ? NUMBER OF SEQUENCES: 1152
 ?

ADDRESS: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
27709

```

1 COMPUTER READABLE FORM:
2 MEDIUM TYPE: Floppy disk
3 OPERATOR: IBM PC compatible
4 OPERATING SYSTEM: PC-DOS/MS-DOS
5 SOFTWARE: PatentIn Release #1.0, Version #1.30
6 CURRENT APPLICATION DATA:
7 APPLICATION NUMBER: US/08/998,416
8 FILING DATE: 24-DEC-1997

```

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: CH 0016/97
 FILING DATE: 31-DEC-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: WEIGER, J. Timothy
 REGISTRATION NUMBER: 38, 241
 REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGL1976
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8587
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 1138:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 719 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: PAG1692UP

US-08-998-416-1138

Query Match	7.3%	Score 37.4	DB 3	Length 719
Sequence Similarity	58.6%	Pred. No. 0.65		
Matches 65	Conservative	0	Mismatches	46
			Indels	0
			Gaps	0

QY 180 TGTCCATTATATTTTATTAGTAGCCGGTTTTCCTACACATCTAAATAGATGAA 239
Db 574 TGAATTATTATTAACTTATTATTGTTTAAACATTATTATTAATGATATAAATATATATAA 515
QY 240 CAGATGCAATGGGTAGTACTGTTTATAAAGAGATATAAAGATCTA 290
Db 514 GGAATATACATTAAATGATATATAGTTTATATAGAACCAAAATGAAGATCTA 464

Search completed: May 27, 2006, 19:25:41
Job time : 107.294 Secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OW nucleic - nucleic search, using sw model

Run on: May 27, 2006, 18:56:27 ; Search time 699.421 Seconds
(without alignments)
8994.972 Million cell updates/sec

Title: US-10-099-663-1_COPY_1115_1626

Perfect score: 512
Sequence: 1 cagtgatccctcattcatcatc.....gtcgtccacagacagaag 512

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:

1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	512	100.0	2381	7 US-10-099-663-1	Sequence 1, Appl
2	326	63.7	413	9 US-10-099-663-2	Sequence 2, Appl
3	43.6	8.5	413	9 US-10-099-663-2	Sequence 2159, Ap
4	42.8	8.4	558	4 US-09-925-065A-572321	Sequence 572321,
5	42.8	8.4	558	4 US-09-925-065A-572321	Sequence 572321,
6	42.8	8.4	679	4 US-09-925-065A-875385	Sequence 875385,
7	42.8	8.4	679	4 US-09-925-065A-875385	Sequence 875385,
8	42.8	8.4	679	4 US-09-925-065A-875386	Sequence 875386,
9	42.8	8.4	679	5 US-09-925-065A-875384	Sequence 875384,
10	42.8	8.4	679	5 US-09-925-065A-875385	Sequence 875385,
11	42.8	8.4	679	5 US-09-925-065A-875386	Sequence 875386,
12	42.8	8.4	900	12 US-10-301-480-582933	Sequence 582933,
13	42.8	8.4	900	12 US-10-301-480-582933	Sequence 1196342,
14	42.6	8.3	1457619	16 US-11-038-686-8739	Sequence 8739, Ap
15	42.4	8.3	5908	7 US-10-221-613-176	Sequence 94, Appl
16	42.4	8.3	5908	7 US-10-221-613-176	Sequence 106, App
17	42.4	8.3	5908	8 US-10-221-613-176	Sequence 176, App

c 18	42.2	8.2	1238	4 US-09-925-065A-710652	Sequence 710652,
c 19	42.2	8.2	1238	5 US-09-925-065A-710652	Sequence 710652,
c 20	41.8	8.2	10048	8 US-10-221-613-211	Sequence 211, App
21	40.8	8.0	536	12 US-10-301-480-220178	Sequence 220178,
22	40.8	8.0	536	12 US-10-301-480-833587	Sequence 833587,
23	40.8	8.0	546	4 US-09-925-065A-121976	Sequence 121976,
24	40.8	8.0	546	5 US-09-925-065A-121976	Sequence 121976,
c 25	40.6	7.9	7040	6 US-10-172-086-13	Sequence 13, Appl
c 26	40.6	7.9	7040	8 US-10-221-714A-161	Sequence 161, App
c 27	40.6	7.9	7040	8 US-10-311-507-47	Sequence 47, Appl
c 28	40.6	7.9	7040	9 US-10-480-846-13	Sequence 13, Appl
c 29	40.6	7.9	7040	9 US-10-473-126-113	Sequence 13, App
c 30	40.6	7.9	7040	9 US-10-473-126-319	Sequence 319, App
c 31	40.6	7.9	14095	7 US-10-311-455-450	Sequence 450, App
c 32	40.4	7.9	470	4 US-09-925-065A-178215	Sequence 178215,
c 33	40.4	7.9	470	4 US-09-925-065A-178215	Sequence 178215,
c 34	40.4	7.9	470	5 US-09-925-065A-178216	Sequence 178216,
c 35	40.4	7.9	470	5 US-09-925-065A-178216	Sequence 178216,
c 36	40.4	7.9	471	12 US-10-301-480-269235	Sequence 269235,
c 37	40.4	7.9	471	12 US-10-301-480-269235	Sequence 269235,
c 38	40.4	7.9	471	12 US-10-301-480-882644	Sequence 882644,
c 39	40.4	7.9	471	12 US-10-301-480-882645	Sequence 882645,
c 40	40.4	7.9	1816	4 US-09-925-065A-45365	Sequence 45365, A
c 41	40.4	7.9	1816	5 US-09-925-065A-45365	Sequence 45365, A
c 42	40.4	7.9	1816	12 US-10-301-480-146603	Sequence 146603,
c 43	40.4	7.9	1816	12 US-10-301-480-146603	Sequence 146603,
c 44	40.2	7.9	1400	16 US-11-136-527-6182	Sequence 6182, Ap
c 45	40.2	7.9	1900	16 US-11-136-527-2086	Sequence 2086, Ap

ALIGNMENTS

RESULT 1
US-10-099-663-1
; Sequence 1, Application US/10099663
; Publication No. US20030177516A1
; GENERAL INFORMATION:
; APPLICANT: Avigenics, Inc
; TITLE OF INVENTION: Avian Gut-Specific Promoters
; FILE REFERENCE: A181
; CURRENT APPLICATION NUMBER: US/10/099,663
; CURRENT FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2381
; TYPE: DNA
; ORGANISM: Gallus gallus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(1626)
; NAME/KEY: exon
; LOCATION: (1627)..(1693)
; NAME/KEY: Intron
; LOCATION: (1694)..(2322)
; NAME/KEY: exon
; LOCATION: (2323)..(2381)
; US-10-099-663-1
Query Match 100.0%; Score 512; DB 7; Length 2381;
Best Local Similarity 100.0%; Pred. No. 1.2e-123;
Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGTGAATCTTCATCTCATGCTTATATGAGATGATGATTCACCAATTGAA 60
DB 1115 CAGTGAATCTTCATCTCATGCTTATATGAGATGATGATTCACCAATTGAA 1174
QY 61 ATGGAACAAAGAGATTTGTTTATCTGTTGGTAAATCGTTTCCAGTTGAT 120
DB 1175 ATGGAACAAAGAGATTTGTTTATCTGTTGGTAAATCGTTTCCAGTTGAT 1234
QY 121 AAGAACCTCCACCAAGTATTAAGTCTATGCAACAAAGAAATGTCATATCTCTT 180

Db 1235 AAAGACCCCTCCACAGATTAAGTCTATGCAAGAAAGATTCATTCCTTT 1294
QY 181 AGTCATTAATTAATTTTCTATTAAGTACCGGTTTTTCTAACAATCAATGATGAC 240
Db 1295 AGTCATTAATTAATTTTCTATTAAGTACCGGTTTTTCTAACAATCAATGATGAC 1354
QY 241 AGAATGAATGGGTTAGTACTGTTTATTAAGAGTAATTAAGTCTATTCATCATTTG 300
Db 1355 AGAATGAATGGGTTAGTACTGTTTATTAAGAGTAATTAAGTCTATTCATCATTTG 1414
QY 301 AGGCAATTAAGGAGGAGAGATTCAGCAACAGTGTGCTTCAAGTGGAAAACAAGTTAA 360
Db 1415 AGGCAATTAAGGAGGAGAGATTCAGCAACAGTGTGCTTCAAGTGGAAAACAAGTTAA 1474
QY 361 ACTAAGTGAACCCCTCTGAGCAAGATCAATGCCACATGAGCTTTAGCCACACACA 420
Db 1475 ACTAAGTGAACCCCTCTGAGCAAGATCAATGCCACATGAGCTTTAGCCACACACA 1534
QY 421 TCATCATGTAATTCCTTCTGATTAAGCCCTGTTCAATAATTCCTTTGCAAAAGCTCTGC 480
Db 1535 TCATCATGTAATTCCTTCTGATTAAGCCCTGTTCAATAATTCCTTTGCAAAAGCTCTGC 1594
QY 481 TACTTACCAAGAGTCTGCTTACAGACAGAAAG 512
Db 1595 TACTTACCAAGAGTCTGCTTACAGACAGAAAG 1626

RESULT 2
US-10-099-663-2
; Sequence 2, Application US/10099663
; Publication No. US20030177516A1
; GENERAL INFORMATION:
; APPLICANT: Avigen, Inc
; TITLE OF INVENTION: Avian Gut-Specific Promoters
; FILE REFERENCE: A181
; CURRENT APPLICATION NUMBER: US/10/099,663
; CURRENT FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Gallus gallus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(336)
; OTHER INFORMATION: chick intestinal fatty acid binding protein promoter region
US-10-099-663-2

Query Match 63.7%; Score 326; DB 7; Length 336;
Best Local Similarity 100.0%; Pred. No. 3.9e-75;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 181 ATTATTAATTTTCTATTAAGTACCGGTTTTTCTAACAATCAATGATGACGAATG 246
Db 1 ATTATTAATTTTCTATTAAGTACCGGTTTTTCTAACAATCAATGATGACGAATG 60
QY 241 AATGAGTTAATGACCTTTAATAAAGATTAATTAAGTCTATTCATCATTTAGGCAA 306
Db 61 AATGAGTTAATGACCTTTAATAAAGATTAATTAAGTCTATTCATCATTTAGGCAA 120
QY 307 TAAAGGAGGAGAGATTCAGCAACAGTGTGCTTCAAGTGGAAAACAAGTTAACTTAA 366
Db 121 TAAAGGAGGAGAGATTCAGCAACAGTGTGCTTCAAGTGGAAAACAAGTTAACTTAA 180
QY 367 GAGACCCCTCTGAGCAAGATCAATGCCACATGAGCTTTAGCCACACATCATCA 426
Db 181 GAGACCCCTCTGAGCAAGATCAATGCCACATGAGCTTTAGCCACACATCATCA 240
QY 427 TGTAAATGCTTCTGATTAAGCCCTGTTCAATAATTCCTTTGCAAAAGCTCTGCTTAA 486
Db 241 TGTAAATGCTTCTGATTAAGCCCTGTTCAATAATTCCTTTGCAAAAGCTCTGCTTAA 300

QY 487 CCAGAGTCTGCTTACAGACAGAAAG 512
Db 301 CCAGAGTCTGCTTACAGACAGAAAG 326

RESULT 3
US-10-674-124A-2159/c
; Sequence 2159, Application US/10674124A
; Publication No. US2004019797A1
; GENERAL INFORMATION:
; APPLICANT: IMOXO, Hideotshi
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; FILE REFERENCE: ORIN-003CIP
; CURRENT APPLICATION NUMBER: US/10/674,124A
; PRIOR FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 2159
; LENGTH: 413
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Located on chromosome 1
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; FEATURE:
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
US-10-674-124A-2159

Query Match 8.5%; Score 43.6; DB 9; Length 413;
Best Local Similarity 55.2%; Pred. No. 0.7;
Matches 85; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 138 TATTAAGTCTTATGCAACAAAGAAATGCAATACATCTCTTATGCTCATTTATTTT 197
Db 256 TATTAAGTCTTATGCAACAAAGAAATGCAATACATCTCTTATGCTCATTTATTTT 197
QY 198 CATTAATTAAGCCGTTTTTCTAACAATCAATGATGACGAATGATGAGTTAGT 257
Db 196 TATTAATTAATTAATTAATGCTCTCACTCAAAAGGCTACCAAGAAAGCAAGTACTAT 137
QY 258 GACTGTTTATTAAGAGCTAATTAAGTACTAT 291
Db 136 GACTGTTTATTAAGAGCTAATTAAGTACTAT 103

RESULT 4
US-09-925-065A-572321/c
; Sequence 572321, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: Nucleotide Polymorphisms in the Human Genome
US-09-925-065A-572321/c

```

; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 572321
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-572321

```

```

Query Match      8.4%; Score 42.8; DB 4; Length 558;
Best Local Similarity 53.6%; Pred. No. 1.3; Indels 0; Gaps 0;
Matches 89; Conservative 0; Mismatches 77;

```

```

QY 187 ATTATTTTTCATTAGATAGCCGGTTTCTTACTCAACTCAATATAGATGAAGCAATG 246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 346 ATTAACATTAAACAATTAGTACCTTTATATCTCTTAACAGACATTTAGAACT 287
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 247 AATGGTTAGTACTGTTTATTAAGAGAGATTAATTAAGTACTATCATTTGAGCAA 306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 286 TAAAGTAAAGTCATGTAATTAAGAAATAGAACTAATTAATTAAGAAATAGCAA 227
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 307 TAAAGAGAGAGAGATTCAGCAACAGTGTCTTACAGTGAATA 352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 226 TAAAGACACTTAATCTTAGAAATTAATTAATAAATGTAACAA 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 5
US-09-925-065A-572321/C
; Sequence 572321, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 572321
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-572321

```

```

Query Match      8.4%; Score 42.8; DB 5; Length 558;
Best Local Similarity 53.6%; Pred. No. 1.3; Indels 0; Gaps 0;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
QY 187 ATTATTTTTCATTAGATAGCCGGTTTCTTACTCAACTCAATATAGATGAAGCAATG 246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 346 ATTAACATTAAACAATTAGTACCTTTATATCTCTTAACAGACATTTAGAACT 287
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 247 AATGGTTAGTACTGTTTATTAAGAGAGATTAATTAAGTACTATCATTTGAGCAA 306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 286 TAAAGTAAAGTCATGTAATTAAGAAATAGAACTAATTAATTAAGAAATAGCAA 227
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 307 TAAAGAGAGAGAGATTTCAGCAACAGTGTCTTCAAGTGAATA 352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 226 TAAAGACACTTAATCTTAGAAATTAATTAATAAATGTAACAA 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 6
US-09-925-065A-875384/C
; Sequence 875384, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 875384
; LENGTH: 679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-875384

```

```

Query Match      8.4%; Score 42.8; DB 4; Length 679;
Best Local Similarity 53.6%; Pred. No. 1.4; Indels 0; Gaps 0;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
QY 187 ATTATTTTTCATTAGATAGCCGGTTTCTTACTCAACTCAATATAGATGAAGCAATG 246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 515 ATTAACATTAAACAATTAGTACCTTTATATCTCTTAACAGACATTTAGAACT 456
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 247 AATGGTTAGTACTGTTTATTAAGAGAGATTAATTAAGTACTATCATTTGAGCAA 306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 455 TAAAGTAAAGTCATGTAATTAAGAAATAGAACTAATTAATTAAGAAATAGCAA 396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 307 TAAAGAGAGAGAGATTTCAGCAACAGTGTCTTCAAGTGAATA 352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 395 TAAAGACACTTAATCTTAGAAATTAATTAATAAATGTAACAA 350
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 7
US-09-925-065A-875385/C
; Sequence 875385, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30

```



```

: TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-875385

```

Query Match	8.4%	Score 42.8;	DB 5;	Length 679;
Best Local Similarity	53.6%;	Pred. No.1.4;		
Matches	89;	Conservative	0;	Mismatches 77;
			Indels	0;
			Gaps	0;

Qy	187	ATTATATATTTTCATTAGATGAGCGGTTTTCCTCAACATCCAATATAGATGAACGATG	246
Db	515	ATTACATTTAACAAATTTAGCTAGCTTTTATATATCTTTTAAACGAACTTTTAGAACT	456
Qy	247	AATGGTTAGTGCCTGTTTATTAAGAGAGATATAAGATCTATCATCTATTGAGCAA	306
Db	455	TAAAGTATAGTCCATGTATTAATTAAGATTAAGATTAAGAACATCTATTAATTAAGATTAAGCAA	356
Qy	307	TAAAGGAGGAGAGATTACGAAACAGTGTGCTTACAAGTGGAAAA	352
Db	395	TAAAGAACCTAATACTTAGAATTAATTAATTAATAATGTACAAA	350

RESULT 11
US-09-925-065A-875386/c
; Sequence 875386, Application US/09925065A
; Publication No. US20050228172A9
JOURNAL TRANSMISSION

```

? TITLE OF INVENTION: Identification and Mapping of Single
? TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
? FILE REFERENCE: 108827.135
? CURRENT APPLICATION NUMBER: US/09/925,065A
? CURRENT FILING DATE: 2001-08-08
? PRIOR APPLICATION NUMBER: US 60/243,096
? PRIOR FILING DATE: 2000-10-24
? PRIOR APPLICATION NUMBER: US 60/252,147
? PRIOR FILING DATE: 2000-11-20
? PRIOR APPLICATION NUMBER: US 60/250,092
? PRIOR FILING DATE: 2000-11-30
? PRIOR APPLICATION NUMBER: US 60/261,766
? PRIOR FILING DATE: 2001-01-16
? PRIOR APPLICATION NUMBER: US 60/289,846
? PRIOR FILING DATE: 2001-05-09
? NUMBER OF SEQ ID NOS: 957086
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 875386
? LENGTH: 679
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-09-925-065A-875386

```

	Query Match	Best Local Similarity	8.4%	Score 42.8	DB 5	Length 679
	Matches	89	Conservative	0	Mismatches	77
					Indels	0
					Gaps	0
QY	187	ATTATTTATTTTCATTAGATAGCCGGTTTTTACTCAACTCAATATAGATGAACAGATG	246			
Db	515	ATTAAACATTAAACAATTAGTAGCCCTTTATATACCTTTAAACAGAACTTTTAGGAAC	456			
QY	247	AATGGGTTAGTGACTGTTTATTAAGAGAATTAATTAAGATCTATCATCTATTGGCCAA	306			
Db	455	TAAAGTAAAGCCATGTATTAATTAAGAAATAGAAATAGAACTATTAATTAAGATTAAGCAA	396			
QY	307	TAAAGGAGGAGAGATTACGAAACAGTGTGCTTACACTGGAAAA	352			
Db	395	TAAAGAACCTAAATTAATCTAGAAATTAATTTAATAAAAATGTACAAA	350			

RESULT 12
US-10-301-480-582933/C
; Sequence 582933, Application US/10301480
; Publication No. US3006005756441
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

```

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 582933
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-582933

Query Match      8.4%  Score 42.8;  DB 12;  Length 900;
Best Local Similarity 53.6%;  Pred. No. 1.6;
Matches      89;  Conservative      0;  Mismatches      77;  Indels      0;  Gaps      0;

```

Oy	187	ATTATTTATTTTCATTAGTAGCCGGTTTTTTACTACAACTCAATTAAGATGAACGAATG	246
Db	485	ATTACATTTAACTTAATTAAGTCTTTTAAATCTCTTTAACAAGCACTTTTAGAAACT	426
Oy	247	AATGGTTAGACGCTTTATTAAGAAGATTAAGTACTATCATGATTTGAGCAA	306
Db	425	TAAAGTAGTAGCCATGTATTAATGAATTAAGAACTATTAATTAAGATTAAGTAA	366
Oy	307	TAAAGGAGGAGAGATTCAGCAAACTGTGCTTACAGTGGAAA	352
Db	365	TAAAGACCTTAATCTTAGAAATTAATTTAATATTAATTTGTAACAA	320

```

RESULT 13
US-10-301-480-1196342/c
; Sequence 1196342, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 122618
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 1196342
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapien
; US-10-301-480-1196342

```

	Query Match	8.4%	Score 42.8	DB 12	Length 900
	Best Local Similarity	53.6%	Pred. No. 1.6		
	Matches	89	Conservative	0	Mismatches 77
				Indels	0
				Gaps	0
QY	187	ATTATTTATTTTCATTAGTACGCGGTTTTTTCACACACTCAAAATTAAGTACACAGAAATG			
Db	485	ATTATACATTAACAAATTAAGTACCTTTATATACCTTTAACACGCACTTTAGGAACT			
QY	247	AATCGGTTAGTGACTGTTTATTAAGCAAGATTAATTAAGATCTCATCATTTGAGCAA			
Db	425	TAAAGATTAAGCCATGTATATATAAAGAAATTAAGAAACACTTAATTAAGATTAAGCAA			
QY	307	TAAAGGAGGAGAGATTCAAGAAACAGTGTCTTCAAGTGGAAAA			
Db	365	TAAAGAACATTAATTAAGTAAATTTATATATAAAATGTACAAA			

```
RESULT 14
US-11-098-686-8739
; Sequence 8739, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8739
; LENGTH: 1457619
; TYPE: DNA
; ORGANISM: Lawsonia intracellularis
US-11-098-686-8739

Query Match      8.3%; Score 42.6; DB 16; Length 1457619;
Best Local Similarity 49.6%; Pred. No. 48;
Matches 167; Conservative 0; Mismatches 164; Indels 6; Gaps 2;

QY 2 AGTGGATCCTTCATCTCATGCTTATTTATGAGATGATAGATTCTGACCAATTTAGAA 61
DB 858309 AGTGATATCTTATGCTTATGATATCACTTAAATAATTAATTCATTAACAGATTTATTA 858368
QY 62 TGGACAAAGCAGAGATTGTGTGTT--TATCTGTGGTAATTAAGCTTTCTCCAGTGT 118
DB 858369 TATATCAAAAAAATTTCTTTTATATTAATGCTACATATTAATAATTAAGTATACCTCTA 858428
QY 119 ATTAAGACCTCCCAAGTATTAAGTCTATGACAAAGAAATGTCATATACATCTC 178
DB 858429 TATATTAACAAATACAAAAAATAAATTAACAAAAATTAATGATTTGCTGTAATTA 858488
QY 179 TTAGCTCATTTATTTTTCATTAGAT--AGCCGTTTTTTTACTACAACTCAAAATTAAGA 235
DB 858489 TTAGCGCTATTAATAGTATATATTAATTAATTAATTAATTAATTAATTAATTAATTA 858548
QY 236 TGAACAGATGAATGGTTAGTACTGTTTATTAAGAAAGATTAATTAATTAATTAATTAATTA 295
DB 858549 TAACTATTAATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 858608
QY 296 ATTTGAGCAATTAAGGAGGAGATTCAGCAACA 332
DB 858609 ATTAAGAGGATTAATTTAGAAAGACTATTAACAAACA 858645

RESULT 15
US-10-239-676-94/C
; Sequence 94, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: Olek, Alexander
; APPLICANT: PIPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
```

```
2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 94
; LENGTH: 5908
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-94

Query Match      8.3%; Score 42.4; DB 6; Length 5908;
Best Local Similarity 54.5%; Pred. No. 4.7;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 138 TATAAGTCCTATGCAACAAGAAATGTCATCTCTAGTCTCATTTATTTATTTT 197
DB 4703 TATACTTCTTTTAAATAAACAATATATATATATATATATATATATATATATATATATAT 4644
QY 198 CATTAGATAGCCGGTTTTTTTACTACAACTCAATATAGATGACAGATGAATGGGTTAGT 257
DB 4643 TATTTATTTTATTTTCTACTACTCTTCTCTAATAATTAATAATTAATAATCTATACCCCTAT 4584
QY 258 GACTGTTTATAAGAGATATTAAGATCTATCA 293
DB 4583 TCCCTTTATCCAAATATATATATCTATATACGAACA 4548

Search completed: May 27, 2006, 20:03:05
Job time : 700.421 secs
```

APPLICANT: HELIX RESEARCH INSTITUTE

THE UNIVERSITY OF CHICAGO


```
/ TITLE OF INVENTION: Novel full length cDNA
/ FILE REFERENCE: H1-A0106
/ CURRENT APPLICATION NUMBER: US/11/293,697
/ CURRENT FILING DATE: 2005-12-05
/ PRIOR APPLICATION NUMBER: US/10/108,260
/ PRIOR FILING DATE: 2002-03-28
/ NUMBER OF SEQ ID NOS: 5458
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 546
/ LENGTH: 1997
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-11-293-697-546
```

```
Query Match      6.4%; Score 33; DB 7; Length 1997;
Best Local Similarity 52.6%; Pred. No. 2.2;
Matches 72; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
```

```
QY 153 AACAAAGAAATGTCATATCTCTTAGTCATATTATTTCATTAGATAGCCGGT 212
DB 1759 AATAGAGATGATGTCATTAATTTCTGTTAAATAATTAAATTCAGATGAGAAAT 1700
QY 213 TTTTACTACAACTCAATTAAGTGAACGAATGAATGGGTTAGTACTGTTTAAAGA 272
DB 1639 TTATCATAAATAATCATATGTATATATCAGATGAATCTTCAGTGTCTTGTGAAAA 1640
QY 273 AGAGTAATAAGATACT 289
DB 1639 CAATATCTACTTACT 1623
```

```
RESULT 3
US-11-293-697-727/c
/ Sequence 727, Application US/11293697
/ Publication No. US20060105376A1
/ GENERAL INFORMATION:
/ APPLICANT: HELIX RESEARCH INSTITUTE
/ TITLE OF INVENTION: Novel full length cDNA
/ FILE REFERENCE: H1-A0106
/ CURRENT APPLICATION NUMBER: US/11/293,697
/ CURRENT FILING DATE: 2005-12-05
/ PRIOR APPLICATION NUMBER: US/10/108,260
/ PRIOR FILING DATE: 2002-03-28
/ NUMBER OF SEQ ID NOS: 5458
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 727
/ LENGTH: 2873
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-11-293-697-727
```

```
Query Match      6.4%; Score 32.8; DB 7; Length 2873;
Best Local Similarity 59.8%; Pred. No. 2.9;
Matches 55; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
```

```
QY 154 ACAAGAAGAAATGTCATATCTCTTAGTCATATTATTTCATTAGATAGCCGGT 213
DB 1468 AAAATTTAATTGAGCTTAACTATTATTGAAAAATTAGGACATTAATTGTGTCT 1409
QY 214 TTTTACTACAACTCAATTAAGTGAACGAAT 245
DB 1408 TTTAATTACCACTAGTTATATGTGAAGAAAT 1377
```

```
RESULT 4
US-11-233-726-22
/ Sequence 22, Application US/11233726
/ Publication No. US20060090216A1
/ GENERAL INFORMATION:
/ APPLICANT: APURA, Nestor
/ APPLICANT: KNOX, Shing
/ APPLICANT: ALEXANDROV, Nikolai
/ APPLICANT: TATARINOVA, Tatiana
```

```
/ APPLICANT: FANG, Yiwen
/ APPLICANT: PENNELL, Roger
/ APPLICANT: LU, Yu-ping
/ APPLICANT: MEDRANO, Leonard
/ APPLICANT: FELDMANN, Kenneth A.
/ APPLICANT: COOK, Zhong C.
/ TITLE OF INVENTION: Promoter, Promoter Control Elements, And Combinations, And Uses Th
/ FILE REFERENCE: 2750-1619PUS2
/ CURRENT APPLICATION NUMBER: US/11/233,726
/ CURRENT FILING DATE: 2005-02-23
/ PRIOR APPLICATION NUMBER: 60/612,891
/ PRIOR FILING DATE: 2004-09-23
/ PRIOR APPLICATION NUMBER: 60/613,134
/ PRIOR FILING DATE: 2004-09-23
/ PRIOR APPLICATION NUMBER: 60/637,174
/ PRIOR FILING DATE: 2004-12-16
/ NUMBER OF SEQ ID NOS: 46
/ SEQ ID NO 22
/ LENGTH: 995
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
```

```
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1009)..(1009)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)..(995)
/ OTHER INFORMATION: Cereb Promoter construct YP0385 as found in Promoter Report #146
US-11-233-726-22
```

```
Query Match      6.4%; Score 32.6; DB 7; Length 995;
Best Local Similarity 55.9%; Pred. No. 2.3;
Matches 62; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
```

```
QY 172 CATTCCTTAGCTCATTAATTATTTTCATTAGATAGCCGGTTTTTACTCAACTCAAT 231
DB 237 CATGCCATATACACAATTTTGTCTGTCGATGATGATGGGTTTATATATTTCTAAA 296
QY 232 AAGATGAACAGATGAATGGGTTAGTGCAGCTTTTAAAGAAGGTATTA 282
DB 297 AAAACTAACATTTTACTGATTTGAGTGAATTAATGACATTAATTA 347
```

```
RESULT 5
US-10-953-349-2269/c
/ Sequence 2269, Application US/10953349
/ Publication No. US20060107345A1
/ GENERAL INFORMATION:
/ APPLICANT: ALEXANDROV, Nikolai et al.
/ TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
/ FILE REFERENCE: 2750-1579PUS2
/ CURRENT APPLICATION NUMBER: US/10/953,349
/ CURRENT FILING DATE: 2004-09-30
/ NUMBER OF SEQ ID NOS: 40252
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 2269
/ LENGTH: 1616
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
US-10-953-349-2269
```

```
Query Match      6.4%; Score 32.6; DB 6; Length 1616;
Best Local Similarity 58.9%; Pred. No. 2.7;
Matches 56; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
```

```
QY 107 TTCCTCAGTTGTATTAAGACCTCCACAGATTAAGTCTTAGCAACAAGAAATGT 166
DB 1598 TTCTTTCTTCTTAATGAAGCATGTACAGTACTATACAGAGAAATCCACAATGAAAA 1539
QY 167 CAATACATCTCTTAGCTCATTAATTATTTTCATT 201
```


Db 1538 CAAAAAAGCTTAAATCTACTTTTCTTTCTTT 1504

RESULT 6

US-10-524-021-12
Sequence 12, Application US/10524021
Publication No. US2006009590A1

GENERAL INFORMATION:

APPLICANT: NAGOYA INDUSTRIAL SCIENCE RESEARCH INSTITUTE
APPLICANT: GIFU INTERNATIONAL INSTITUTE OF BIOTECHNOLOGY

APPLICANT: YAMADA, Yoshiji

APPLICANT: YOKOTA, Mitsuhito

TITLE OF INVENTION: Method for diagnosing a risk of restenosis after percutaneous cor

FILE REFERENCE: C0200501

CURRENT APPLICATION NUMBER: US/10/524,021

PRIOR FILING DATE: 2005-02-09

PRIOR APPLICATION NUMBER: JP P2002-233041

NUMBER OF SEQ ID NOS: 67

SOFTWARE: PatentIn version 3.1

SEQ ID NO 12

LENGTH: 5204

TYPE: DNA

ORGANISM: Homo sapiens

US-10-524-021-12

Query Match 6.4%; Score 32.6; DB 6; Length 5204;

Best Local Similarity 58.9%; Pred. No. 4;

Matches 56; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Db 138 TATAAAGTCTTATGCAACAAAGAAATGTCATATCTCTAGTCTGATTATTATTT 197

Db 2987 TAGAAGCATCATTTTCAACAAATTTTAAACCTTTTATAGCTCAACATTTTCT 3046

Qy 198 CATTGATGCGCGTTTCTTACTACACTCAATA 232

Db 3047 ATTAAATTACATGTTTGTAATGACATTTTA 3081

RESULT 7
US-11-217-529-76807/c
Sequence 76807, Application US/11217529
Publication No. US2006009612A1

GENERAL INFORMATION:

APPLICANT: SUNTORY LIMITED

APPLICANT: NAKAO, YOSHIHIRO

APPLICANT: NAKAMURA, NORIHISA

APPLICANT: KODAMA, YUKIKO

APPLICANT: FUJIMURA, TOMOKO

APPLICANT: ASHIKARI, TOSHIHIRO

TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS

FILE REFERENCE: S-38-285

CURRENT APPLICATION NUMBER: US/11/217,529

PRIOR FILING DATE: 2005-09-02

PRIOR APPLICATION NUMBER: US 10/932,182

NUMBER OF SEQ ID NOS: 197023

SOFTWARE: PatentIn version 3.3

SEQ ID NO 76807

LENGTH: 1221

TYPE: DNA

ORGANISM: Saccharomyces pastorianus

US-11-217-529-76807

Query Match 6.3%; Score 32.4; DB 7; Length 1221;

Best Local Similarity 57.2%; Pred. No. 2.8;

Matches 79; Conservative 0; Mismatches 56; Indels 3; Gaps 1;

Qy 168 AATATATCTCTACTCATATATTTTCTTATGATGCGGTTTCTTACTACATC 227

Db 601 AATATATCTCTTAACTCATATTAATGCGCATGGGTCAATTTGTTCAATTCCTTTTCT 542

Qy 228 AATATAGATGAACAGATGATGGTTAGTACTGTTAT---AAAGAGATATAAG 284

Db 541 ACTTCTCAAGACAGATATCAAGCGCTTACAGTGTACTCTATATAACAGATATG 482

Qy 285 ATACTATCATCTTTGAG 302

Db 481 ATCTATCATTAATTTTAA 464

RESULT 8

US-10-953-349-24676

Sequence 24676, Application US/10953349

Publication No. US20060107345A1

GENERAL INFORMATION:

APPLICANT: ALEXANDROV, Nikolai et al.

TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES

FILE REFERENCE: 2750-1579PUS2

CURRENT APPLICATION NUMBER: US/10/953,349

PRIOR FILING DATE: 2004-09-30

NUMBER OF SEQ ID NOS: 40252

SOFTWARE: PatentIn version 3.3

SEQ ID NO 24676

LENGTH: 3064

TYPE: DNA

ORGANISM: Glycine max

US-10-953-349-24676

Query Match 6.3%; Score 32.4; DB 6; Length 3064;

Best Local Similarity 50.6%; Pred. No. 3.8;

Matches 78; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 219 CTCAACTCAATTAAGATGAACAGATGATGGTTAGTACTGTTATATAAGAGATA 278

Db 2841 CTAGTAGGCGCATGAAATGAAATTAATAGTAATTTATTTGGCTTTGTAAGAGAA 2900

Qy 279 ATTAAGATATCTATCATCTTTGAGGCAATTAAGGAGAGAGATTCAGCAACATGTGC 338

Db 2901 AAAAAGAAAGAAAAAACAACGTTGATTTAAAAATTTCTAGCTCACCGATGGGGGTG 2960

Qy 339 TTCAAGTGAAGAAACAAGTTAACTTAAGTGACC 372

Db 2961 TTGAACCTGAAATGATTTGTGATTAAGATCTCC 2994

RESULT 9

US-11-301-554-1289

Sequence 1289, Application US/11301554

Publication No. US20060088527A1

GENERAL INFORMATION:

APPLICANT: Henderson, Robert A.

APPLICANT: Wang, Tonglong

APPLICANT: Watanabe, Yoshihiro

APPLICANT: Kalos, Michael D.

APPLICANT: Sleath, Paul R.

APPLICANT: Johnson, Jeffrey C.

APPLICANT: Retter, Marc W.

APPLICANT: Durham, Margarita

APPLICANT: Carter, Derrick

APPLICANT: Fanger, Gary R.

APPLICANT: Vedivick, Thomas S.

APPLICANT: Bangur, Chaitanya S.

APPLICANT: McNabb, Andria

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

FILE REFERENCE: AND DIAGNOSIS OF LONG CANCER

CURRENT APPLICATION NUMBER: US/11/301,554

PRIOR FILING DATE: 2005-12-13

PRIOR APPLICATION NUMBER: US 10/283,017

PRIOR FILING DATE: 2002-10-28

PRIOR APPLICATION NUMBER: US 10/113,872

PRIOR FILING DATE: 2002-03-28

PRIOR APPLICATION NUMBER: US 10/017,754

```

? PRIOR FILING DATE: 2001-10-29
? PRIOR APPLICATION NUMBER: US 09/502,941
? PRIOR FILING DATE: 2001-07-10
? PRIOR APPLICATION NUMBER: US 09/849,626
? PRIOR FILING DATE: 2001-05-03
? PRIOR APPLICATION NUMBER: US 09/736,457
? PRIOR FILING DATE: 2000-12-13
? PRIOR APPLICATION NUMBER: US 09/702,705
? PRIOR FILING DATE: 2000-10-30
? PRIOR APPLICATION NUMBER: US 09/677,419
? PRIOR FILING DATE: 2000-10-06
? PRIOR APPLICATION NUMBER: US 09/671,325
? PRIOR FILING DATE: 2000-09-26
? PRIOR APPLICATION NUMBER: US 09/658,824
? Remaining Prior Application data removed - See File Wrapper or PALM.
? NUMBER OF SEQ ID NOS: 2157
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 1289
?
? LENGTH: 670
?
? TYPE: DNA
? ORGANISM: Homo sapiens
US-11-301-554-1289

```

```

Query Match      6.2%; Score 32; DB 7; Length 670;
Best Local Similarity 45.8%; Pred. No. 2.9;
Matches 110; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY      TCTCCACCAATTAGAATGACAAAGCAGAGATTGCTTTTATCTGTGGTAAATACGT 105
Db       TTTCAATTTAAATATAGAGGTGAATGTTAAATACCTGTTATACATGTTGAATTCATTTAT 167
QY      106 TTTCTCCAGTTGTATTAAGACCCCTCCACCAAGTATAAGTCTCATGCAACAAAGAAATG 165
Db       CTGAAAAATGTATATAAAAAACACACATGTAAGCTGATTTGAGGAGAAAGAAAAATTCATT 227
QY      166 TCAATACATTTCTCTTAGTCTCATTTATATTTCCTATTAGATAGCCGGTTTTTATCTCAAC 225
Db       TTTGTAAATTTTCCCAAGGTTTAAAGATTTTACCAAGAACTTATTCATAGTTTAAAGATGCAA 287
QY      226 TCAATAATAGATGAACAGATGAATGGGTTAGACGCTTTATTAAGAAGATTAATAAGA 285
Db       TTAGGTTCGAAACTTTCAAAGAAAGGGGTGTAGGTGATTAATGAACAGTCACTTAACA 347

RESULT 10
US-10-953-349-3916/c
; Sequence 3916, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3916
; LENGTH: 2276
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-3916

```

```
Db                               ||| |||| | ||| | ||| |  
      2104 TTGATTGTAGTACTAGATGGTAATGAATGAACT 2071  
  
RESULT 11  
US-10-505-928-607            ; Sequence 607, Application US/10505928  
; Publication No. US20060088532A1  
GENERAL INFORMATION:  
APPLICANT: Ludwig Institute for Cancer Research et al.  
TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES  
FILE REFERENCE: 28967/39178  
CURRENT APPLICATION NUMBER: US/10/505,928  
CURRENT FILING DATE: 2004-08-27  
PRIOR APPLICATION NUMBER: US 60/363,019  
PRIOR FILING DATE: 2002-03-07  
NUMBER OF SEQ ID NOS: 866  
SOFTWARE: PatentIn 3.2  
SEQ ID NO 607  
LENGTH: 135090  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-505-928-607
```

```

Query Match          6.1%; Score 31.2; DB 6; Length 135090;
Best Local Similarity 51.4%; Pred. No. 30;
Matches 72; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy      113 AGTGTATAAGACCCGCCACGATATAAGTCCTATGCACAAGAANAATGCAATAC 172
         ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      19941 AGATGTTTAATGTTCAACCCTCCTGTBAGCAGATTATTAATAAATAATAATATTTTG 20000

Qy      173 ATTCTCTTACTCTCATTTATTTTTCATTAGATAGCCGGTTTTTTTACTCAACTCAATA 232
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      20001 ATTTTGCTCTTCTCATTTTCTCTTATTAATAAAAAAACTTGGAAGAACCCCATCTTACTA 20060

Qy      233 AGATGAACAGAAATGATGGG 252
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      20061 AAATACAAAATTAGCTGGG 20080

RESULT 12
US-11-217-529-81653
Sequence 81653, Application US/11217529
Publication No. US20060099612A1
GENERAL INFORMATION:
APPLICANT: SUNTOY LIMITED
APPLICANT: NAKAO YOSHITIRO
APPLICANT: NAKAMURA NORIHISA
APPLICANT: KODAMA YUKIKO
APPLICANT: FUJIMURA TOMOKO
APPLICANT: ASHIMIZARI TOSHIHIRO
TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
FILE REFERENCE: S-38-285
CURRENT APPLICATION NUMBER: US/11/217,529
CURRENT FILING DATE: 2005-09-02
PRIOR APPLICATION NUMBER: US 10/932,182
PRIORITY FILING DATE: 2004-09-02
NUMBER OF SEQ ID NOS: 197023
SOFTWARE: PatentIn version 3.3
SEQ ID NO 81653
LENGTH: 600
TYPE: DNA
ORGANISM: Saccharomyces pastorianus
US-11-217-529-81653
```

Query Match	6.2%;	Score 31.6;	DB 6;	Length 2276;
Best Local Similarity	58.5%;	Prod No. 5.7;		
Matches 55;	Conservative 0;	Mismatches 39;	Indels 0;	Gaps 0;
QY	196	TTCATTAGATGACCGCGTTT	TACTCAACTCAATATAGATGAACAGATCAATGGGTTA	255
DB	2164	TTCCAAAGAAACACGCTTCTCAATATAAAAACAAAAACAACAGCATGATGATGAAGACA		2105
Y	256	GTGACTGTTTATTAAGAAGAGTATATAAAGTACT		289

Query Match	6.1%;	Score 31;	DB 7;	Length 600;
Best Local Similarity	59.8%;	Pred. No. 5.3;		
Matches	52;	Conservative	0;	Mismatches 35; Indels 0; Gaps 0;
Oy	169	ATACATTCCTTAGTCACTATATATATTTTCATATAGATCCGGTTTTTACTACAACTCA	228	
b	125	ATACCTTATCTGGTCACATTTCTTTATCTTTTCAAGAAATATATTTTCAACACATACA	184	

THIS PAGE BLANK (USPTO)